NMR studies of the solution conformation of the sex peptide from Drosophila melanogaster

Moehle, K ; Freund, A ; Kubli, E ; Robinson, J A

Abstract: The insect sex peptide (SP) elicits a variety of biological responses upon transfer to the mated female. SP contains 36 amino acids, including a tryptophan-rich N-terminal region, a central region containing five hydroxyproline (Hyp) residues, and a C-terminal region enclosed by a disulfide bridge. The solution structure of SP, studied here using NMR spectroscopy, includes a motif WPWN that adopts a type I -turn in the N-terminal Trp-rich region. This turn region is connected to the central Hyp-rich region, which adopts extended and/or PPII-like conformations. The C-terminal disulfide-bonded loop populates helical turns or nascent helical structure. Overall, the results reveal a rather flexible peptide that lacks a compact folded structure in solution.

DOI: https://doi.org/10.1016/j.febslet.2011.03.040

NMR Studies of the Solution Conformation of the Sex Peptide from *Drosophila melanogaster*

Kerstin Moehle[a], Annabelle Freund[a], Eric Kubli[b] and John A. Robinson[a]*

[a] Chemistry Department, University of Zurich, Winterthurerstrasse 190, 8057 Zurich, Switzerland,

[b] Institute of Molecular Life Sciences, University of Zurich, Winterthurerstrasse 190, 8057 Zurich, Switzerland.

* Corresponding author

E-mail: robinson@oci.uzh.ch

Abstract

The insect sex peptide (SP) elicits a variety of biological responses upon transfer to the mated female. SP contains 36 amino acids, including a tryptophan-rich N-terminal region, a central region containing five hydroxyproline (Hyp) residues, and a C-terminal region enclosed by a disulfide bridge. The solution structure of SP, studied here using NMR spectroscopy, includes a motif WPWN that adopts a type I ß-turn in the N-terminal Trp-rich region. This turn region is connected to the central Hyp-rich region, which adopts extended and/or PPII-like conformations. The C-terminal disulfide-bonded loop populates helical turns or nascent helical structure. Overall, the results reveal a rather flexible peptide that lacks a compact folded structure in solution.

Keywords:

peptide, secondary structure, conformation, sex peptide, NMR spectroscopy, fruit fly, drosophila

1. Introduction

The peptide hormone known as sex peptide (SP) is produced in seminal fluid of the male fruit fly (*Drosophila melanogaster*), where it elicits a variety of responses upon transfer to the mated female. These include changes in female sexual behaviour (oviposition and receptivity), as well as food intake, egg production, Juvenile hormone (JH) synthesis, ovulation, and antimicrobial peptide synthesis [1,2]. Thus SP probably interacts with several different receptors. One of these receptors was shown recently to be a G-protein coupled receptor (GPCR) called the sex peptide receptor (SPR) or CG16752 [3], which is located in various parts of the insect neuronal and ovarian systems [4]. The interaction of SP with SPR

was shown to be responsible for eliciting oviposition and reduction of receptivity [3]. A different insect GPCR called Methuselah, which is associated with longevity, is also activated by binding to SP but the physiological relevance of this interaction remains unclear [5]. In addition, SP has been shown to elicit a systemic and epithelial innate immune response involving the production of the insect antimicrobial peptide drosocin, suggesting that SP may interact with pattern recognition receptors in the innate immune system of the insect [6,7].

The secreted mature SP contains 36 amino acids, including a tryptophan-rich N-terminal segment, a central region containing five hydroxyproline (Hyp) residues, and a C-terminal region enclosed by a disulfide bridge (Figure-1) [8,9]. The Hyp was shown to be (2*S*,4*R*)-4-hydroxyproline by comparison to a synthetic standard [9], and is consistent with the ability of all known prolyl-4-hydroxylases, including that from *D. melanogaster*, to incorporate the hydroxyl group in the *R* configuration. The biological functions of SP appear to be mediated by specific parts of the peptide [6]. Thus SPR interacts with the C-terminal disulfide-bridged loop, as this is the region needed for binding of SP to the nervous system and genital tract of females, and the same region is also essential for eliciting oviposition and reduction of receptivity [10,11]. The amino-terminal portion of SP is essential for inducing the synthesis of JH and for binding to sperm [12], whereas the Hyp-rich central region is an elicitor of the innate immune response, perhaps acting by chemical mimicry of sugar components of the bacterial cell wall [6].

In this work we investigate the conformational behavior of SP using NMR spectroscopy in water and water/trifluorethanol (TFE) solution. Information about the structure of SP in solution is of interest in understanding how this peptide hormone might interact with various different receptors in the insect nervous, ovarian and immune systems.

2. Material and Methods

2.1. Peptide synthesis

SP was synthesized on an *Applied Biosystems* peptide synthesizer ABI 433A using standard Fmoc chemistry. The following side-chain protected amino acids were used for the synthesis: Fmoc-Ala-OH, Fmoc-Asn(Trt)-OH, Fmoc-Asp(tBu)-OH, Fmoc-Arg(Pbf)-OH, Fmoc-Cys(Trt)-OH, Fmoc-Glu(tBu)-OH, Fmoc-Gly-OH, Fmoc-Ile-OH, Fmoc-Leu-OH, Fmoc-Lys(Boc)-OH, Fmoc-Phe-OH, Fmoc-Hyp(tBu)-OH, Fmoc-Ser(tBu)-OH, Fmoc-Thr(tBu)-OH and Fmoc-Trp(Boc)-OH. Fmoc deprotection was performed using 20% piperidine in NMP. Coupling reactions were carried out using HBTU/HOBt (4 eq.) in DMF and DIEA (8 eq.) in NMP. The activated amino acids (4 eq.) were coupled for 30 min. Capping was performed at the end of each cycle using a solution of HOBt (0.015 M), Ac_2O (0.5 M) and DIEA (0.125 M) in NMP.

SP was assembled on 2-chlorotrityl resin preloaded with Fmoc-Cys(Trt)-OH (0.24 mmol, 0.4 mmol/g, 600 mg). At the end of the synthesis the peptide was cleaved from the resin and deprotected in a solution of TFA:EDT:H_2O:TIS 92.5:2.5:2.5:2.5 (30 mL) with shaking for 4 h at room temp. After concentration *in vacuo*, the crude peptide was precipitated by addition of cold Et_2O (10 mL), collected by centrifugation, washed twice with Et_2O and dried *in vacuo*. The crude product was dissolved in H_2O + 1% TFA (15 mL) and the solution was stored overnight at 4°C to complete the deprotection of tryptophan residues. The crude peptide was then directly purified by reverse-phase HPLC on a preparative C_{18} column (Zorbax Eclipse XDB-C18; 21 mm x 250 mm, 10μm, 80 Å) using a gradient of 15% to 50% acetonitrile in water (+ 0.1% TFA) over 16 min. Reduced SP was obtained as a white powder (90 mg, 20.4 μmol, 17% yield). HPLC purity >98% on an analytical C_{18} column (Zorbax Eclipse XDB-C18; 4.6 mm x 250 mm, 5 μm, 80 Å) using a gradient of 5% to 50% acetonitrile in water (+ 0.1% TFA) over 16 min. R_t = 15.7 min. MALDI-TOF: (*m/z*) 4429.1 $[M+H]^+$ (± 0.1%; $M_{calc.}$ = 4428.1).

The reduced peptide (4.5 μmol, 20 mg) was dissolved in a solution of ammonium acetate buffer (20 ml, 50 mM, pH 8), and stirred in air at room temp. for 40 h. The peptide was purified by reverse-phase HPLC on a semi-preparative C_{18} column (Zorbax Eclipse XDB-C18; 10 mm x 250 mm, 10μm, 80 Å) using a gradient of 5% to 50% acetonitrile in water (+ 0.1% TFA) over 16 min. After lyophilization SP was obtained as a white solid (12 mg) in 60% yield. The product was ≥98% pure by reverse-phase HPLC (analytical C_{18} column, Zorbax Eclipse XDB-C18; 4.6 mm x 250 mm, 5 μm, 80 Å, using a gradient of 5% to 50% acetonitrile in water (+ 0.1% TFA) over 16 min). R_t = 15.7 min. MALDI-TOF: (*m/z*) 4427.3 $[M+H]^+$ (± 0.1%; $M_{calc.}$= 4426.3).

Synthetic SP was assayed for oviposition and receptivity according to standard procedures [8], and was found to show full activity.

2.2 NMR spectroscopy and structure calculations

^{1}H NMR measurements were performed in H_2O/D_2O (9:1) or pure D_2O, pH 5.0; or in a mixture (1:1) of trifluoroethanol (TFE-d_3) and H_2O/D_2O (9:1), pH 5.0. Spectra were acquired on either a Bruker AV-600 or AV-700 spectrometer at 300 K. Data were processed using TOPSPIN 2.1 (Bruker) or XEASY [13]. 1D ^{1}H NMR spectra were also recorded at various temperatures (279 K, 286 K, 293 K, 307 K and 314 K), pH values (4, 6 and 7) and SP concentrations (0.07 mM, 0.13 mM, 0.27 mM and 0.68 mM). Water suppression was performed by presaturation. Spectral assignments were made for water and TFE/water solutions using 2D DQF-COSY, TOCSY and NOESY spectra. $^3J_{HN\alpha}$ coupling constants were determined from 1D spectra or from 2D NOESY spectra by inverse Fourier transformation of in-phase multiplets [14].

[^{15}N, ^{1}H]- and [^{13}C,^{1}H]-HSQC spectra at natural abundance were recorded using 3 mM peptide samples in H_2O/D_2O 9:1, and in pure D_2O. Acquisition data were sampled as

1024 x 128 or 1024 x 200 complex points in t2 and t1, for the ^{15}N and ^{13}C correlation spectra, respectively. The 2D data matrices were multiplied by a square-sine-bell window function and zero filled to 2k x 1k or 2k x 2k points prior to Fourier transformation. Baseline correction was applied in both dimensions. Spectra were referenced in the direct dimension to the water signal and in the ^{15}N dimension using the ^{15}N to ^{1}H frequency ratio of 0.101329118, while carbon referencing was directly performed on the internal TSP standard (0 ppm).

Distance restraints were obtained from NOESY spectra with a mixing time of 200 ms. Spectra were typically collected with 1024 x 256 complex data points zero-filled prior to Fourier transformation to 2048 x 1024, and transformed with a cosine-bell weighting function. The structure calculations were performed by restrained molecular dynamics in torsion angle space using DYANA [15], using only the NOE-derived distance restraints. Starting from 100 randomized conformations a bundle of 20 conformations were selected, which encur the lowest DYANA target energy function. The program MOLMOL [16] was used for structure analysis and visualization of the molecular models. The structures have been deposited in the Protein Data Base.

Time-averaged restrained MD simulations were carried out using GROMACS v4.5.3 [17] with the AMBER99 force field in explicit solvent and periodic boundary conditions. The initial structure was selected from the calculated NMR structures. The peptide was placed in a rhombic dodecahedral unit cell with a minimum distance of 1.2 nm to the box edge. TIP3P water molecules and ions to counterbalance the peptide charge were added yielding a final system of 16338 water molecules and 5 Cl^- ions. Steepest decent minimization was performed. Electrostatics were treated with particle-mesh Ewald (PME) using a short range cut-off, and a cut-off for van der Waals forces of 1.4 nm. After 20 ps equilibration with position restraints on peptide heavy atoms, a 10 ns simulation with time-averaged NOE

distance restraints (Table-S3) was performed. In all simulations the temperature was maintained close to 300 K by weak coupling to an external bath with a coupling constant of 0.1 ps. The LINCS algorithm was used to constrain all bond lengths.

3. Results and Discussion

A series of 1D ^{1}H NMR spectra at varying pH and peptide concentrations revealed only marginal changes in shape and dispersion of NMR signals, suggesting that the peptide does not aggregate significantly under these conditions. ^{1}H NMR spectra in water and in TFE/water (1:1) mixtures revealed two species in a ratio of approximately 1:3 and 1:5, respectively, which interconvert slowly on the NMR timescale, representing *cis-trans* isomers at the Trp^{3}-Pro^{4} peptide bond. The major species was the *trans* isomer at all Xaa-Pro and Xaa-Hyp bonds, as indicated by Hα-Hδ NOE connectivities between the corresponding residues. When NMR spectra were recorded in D_2O, all the peptide NH groups exchanged rapidly (within the time of the NMR measurement), indicating the absence of stable intramolecular hydrogen bonded structures.

Chemical shifts. ^{1}H and ^{13}C chemical shift deviations (CSDs) from random coil values ($\Delta\delta=\delta^{observed} - \delta^{random}$) were examined for indications of regular secondary structure (Figure-2) [18]. The patterns expected are positive values for $\Delta\delta_{H_\alpha}$ and $\Delta\delta_{C\beta}$ and negative values for $\Delta\delta_{C_\alpha}$ in ß-strands, whereas values of opposite sign are expected for α-helices [19]. The CSDs in the N-terminal Trp-rich region should be interpreted with caution, since they may be influenced by aromatic ring current as well as conformational effects [20], although analyses of NOEs provide evidence for a ß-turn at the Trp^{3}Pro^{4}Trp^{5}Asn^{6} motif (see below). In the central Hyp-rich region (the CSDs of Hyp have not been documented), the $\Delta\delta_{H_\alpha}$ and $\Delta\delta_{C_\alpha}$ values for non-Hyp residues clearly suggest that ß-structure is favored (Figure-2). The numerous Hyp residues, however, may have a special effect on the conformation of the

peptide backbone in the central region. For example, the role of Hyp in stabilizing the triple helix in collagen is well known [21]. Also, proline residues are generally restricted to two narrow regions of dihedral space: the α-helical region at (ϕ,ψ) = (-61°,-35°) and the β-region at (-65°, +150°) [22,23]. The extended conformation – the so-called left-handed polyproline II helical conformation (PPII) – is the dominant fold in pure polyproline peptides. The PPII conformation is an extended helical structure, with three residues per turn, which does not form intramolecular backbone hydrogen bonds [23]. While the PPII helix has a preference for proline residues, it is also found with other amino acids [23]. Furthermore, proline restricts the preceding residue to the β-region of dihedral (ϕ,ψ) space [22]. The CSDs observed for each amino acid preceding a Hyp residue are consistent with β- and/or PPII conformations. In the C-terminal region from Arg^{20}-Asn^{27}, the negative $\Delta\delta_{H\alpha}$, positive $\Delta\delta_{C\alpha}$, and negative $\Delta\delta_{C\beta}$ values provide support here for nascent helical secondary structure (Figure-2). From Cys^{24} this includes several residues that form part of the disulfide-bridged loop. From Leu^{28} to the C-terminus this trend into negative CSDs is less apparent, suggesting that nascent helical structure does not extend into the second half of the disulfide-bridged loop. Interestingly, at the end of the helical section (Arg^{20}-Asn^{27}) the motif $Gly^{29}Pro^{30}Ala^{31}$ possesses a high propensity for ß-turn formation. Indeed, NOEs are observed that support the occurrence of a nascent turn in this region of the SP backbone (see below).

All measured chemical shifts, including HA, CA, CB, HN and N (see Supporting information), were also analyzed using TALOS+, a program that exploits the large database of protein NMR data to establish accurate relations between chemical shifts and backbone torsion angles ϕ and ψ [24]. Using these data, TALOS+ predicts with a "good" score (RCI-derived order parameter S^2 from 0.69-0.77) that all non-Hyp residues from Phe^{12}-Asn^{18} will occupy the ß-region, and all residues from Trp^{23}-Leu^{28} the α-region of ϕ/ψ space.

Coupling constants. Backbone $^{3}J_{HN\alpha}$ coupling constants ≤5 Hz are typical for residues in α-helices and >7Hz for residues in regular ß-structure [25]. However, most of the $^{3}J_{HN\alpha}$ values for SP in water are in the range 6-7 Hz (Figure 3), which is consistent with a relatively flexible backbone and/or with PPII conformations. However, a series of lower $^{3}J_{HN\alpha}$ values are seen in the region Arg^{20}-Asn^{27} in water/TFE, adding further support for nascent helical structures in this region.

Nuclear Overhauser Effects. 2D-NOESY spectra revealed few long range NOEs (see below), suggesting that SP does not adopt a compact folded structure under the conditions tested (Figure-3). Average solution structures for SP were calculated with DYANA using NOE-derived distance restraints. Rather high mean rmsd values are obtained after global backbone superimposition of the final 20 structures. The observed NOEs, however, provide evidence for ordered secondary structure, or nascent structure, only in distinct regions of SP, which is supported also by analyses of CSDs and coupling constants (see above). The three regions of interest are the N-terminal Trp-rich sequence (Glu^{2}-Lys^{8}), the Hyp-rich central region (Phe^{12}-Asn^{18}), and the disulfide-bonded loop region (Lys^{22}-Cys^{36}). Upon superimposition of these individual regions in the final 20 structures, lower backbone rmsd values are obtained (Table-1). In addition, time-averaged NOE-derived distance restraints were applied in MD simulations in order to take account of flexibility in the SP backbone (see Supporting information, Figure S2).

In the N-terminal Trp-rich region of SP in TFE/water, sequential as well as multiple medium range NOEs (between Trp^{3} and Trp^{5}/Asn^{6} - see Supporting information for a full list of NOE distance restraints) provide strong evidence for a well-defined type I ß-turn in the WPWN motif. This turn is highly populated in the NMR structures calculated using DYANA (Figure 4). This turn occurs in the *trans*-Trp^{3}-Pro^{4} rotamer, and is characterized by the Trp^{5} side chain stacking against the pyrrolidine ring of Pro^{4} (Figure-4B and C) and the side-chain

of Trp^3. A much stronger Pro^4 Hα to Trp^5 HN NOE than actually observed would be expected if a type II turn were formed. Evidence for a type I turn is the upfield shifted Hα resonance for Pro^4 and downfield shifted of Hα in Trp^5 (Figure 2), a combination that is typical for position 2/3 of only the type I turn [26]. On the other hand, for SP dissolved in water only two weak medium range NOEs are seen between Trp^3 and Trp^5. This suggests that a nascent rather than a stable turn is populated in this region in pure water. During the course of 10 ns time-averaged distance-restrained MD simulations (TAMD) in explicit water solvent, the Pro^4-Trp^5 remained in a type I turn (Figure S2).

In the hydroxyproline-rich central region (Hyp^9-Hyp^{19}), only intra-residue and sequential NOEs are observed in both solvents. This and the observation of strong Hα(*i*)-Hδ(*i*+1) and weak or absent HN(*i*)-Hδ(*i*+1) NOEs, for each Xaa-Hyp pair, are consistent with extended and/or PPII conformations in this region. The same conclusion was also suggested from the CSD data, discussed above. It seems that this region may be extended and semi-rigid, as seen in the NMR structures shown in Figure 4. The TAMD simulations support the high propensity of PPII conformations in this region, although turn conformations were also observed at Ser^{16}-Hyp^{17} , and Hyp^{19}-Arg^{20} (see Supporting information).

In the region of the disulfide-bonded loop (Arg^{20}-Cys^{36}), a series of sequential HN-HN as well as medium range Hα(*i*)-NH(*i*+2) and Hα(*i*)-NH(*i*+3) NOEs are seen from Arg^{20} to Asn^{27}, particularly in water/TFE mixture. These NOEs are characteristic of α-helical conformations, or a nascent helix (in water) in this region, as suggested also by the CSD data discussed above. However, the TAMD simulations (Figure-S2) suggest more flexibility may be present in this region. A reversal of the peptide chain direction is needed for disulfide ring closure at the C-terminus, which is achieved by a turn conformation at Pro^{30}-Ala^{31}. A strong HN-HN NOE between Trp^{32}-Ala^{31}, a Hα–HN NOE between Trp^{32}-Pro^{30}, as well as the upfield shifted HN of Trp^{32} also support the occurrence here of turn-like structures. A type-I

ß-turn conformation at Pro^{30}-Ala^{31} is also stably maintained in the TAMD trajectory (Figure S2), in agreement with the NMR data. Furthermore, long range NOEs between the side chains of Leu^{26}-Trp^{23}, and between Trp^{23}, Leu^{28} and Trp^{32} indicate that these amino acid side chains are clustered into a hydrophobic patch on the same side of the macrocycle (Figure 4). Consistent with this clustering is the upfield shift of the Asn^{27} side chain protons, which in DYANA structures come into close contact to the Trp^{23} aromatic side chain. However, no evidence of regular secondary structure was observed in the final C-terminal segment between Trp^{32} and Cys^{36}.

The overall picture to emerge from these studies is of a flexible peptide that lacks a compact folded structure, but which nevertheless possesses elements of nascent secondary structure in distinct regions, in aqueous solution. This nascent structure becomes more pronounced in water/TFE mixtures. The N-terminal Trp-rich region contains a motif WPWN that forms type I ß-turns. This turn region is connected to a central Hyp-rich region, which populates an extended and/or PPII-like conformation. Finally, towards the C-terminus and the disulfide-bonded loop, a region of nascent helical and another ß-turn is found. These distinct regions of nascent secondary structure may be important for recognition of SP by its various cellular receptors.

Acknowledgements

This work was supported by the University of Zurich. Prof. Oliver Zerbe is thanked for help with measurement of heteronuclear NMR spectra.

References

[1] Kubli, E. (2008) Curr. Biol. 18, R210-2.

[2] Kubli, E. (2010) Curr. Biol. 20, R474-6.

[3] Yapici, N., Kim, Y.-J., Ribeiro, C. and Dickson, B.J. (2007) Nature 451, 33-37.

[4] Häsemeyer, M., Yapici, N., Heberlein, U. and Dickson, B.J. (2009) Neuron 61, 511-518.

[5] Ja, W.W., Carvalho, G.B., Madrigal, M., Roberts, R.W. and Benzer, S. (2009) Protein Sci. 18, 2203-2208.

[6] Domanitskaya, E.V., Liu, H.F., Chen, S.J. and Kubli, E. (2007) FEBS Journal 274, 5659-5668.

[7] Peng, J., Zipperlen, P. and Kubli, E. (2005) Curr. Biol. 15, 1690-1694.

[8] Chen, P.S., Stumm-Zollinger, E., Aigaki, T., Balmer, J., Bienz, M. and Böhlen, P. (1988) Cell 54, 291-298.

[9] Schmidt, T., Choffat, Y., Klauser, S. and Kubli, E. (1993) J. Insect Physiol. 39, 361-368.

[10] Ding, Z., Haussmann, I., Ottiger, M. and Kubli, E. (2003) J. Neurobiol. 55, 372-384.

[11] Kubli, E. (2003) Cell Mol. Life Sci. 60, 1689-1704.

[12] Peng, J., Chen, S., B¸sser, S., Liu, H., Honegger, T. and Kubli, E. (2005) Curr. Biol. 15, 207-213.

[13] Bartels, C., Xia, T.-h., Billeter, M., Güntert, P. and Wüthrich, K. (1995) J. Biomol. NMR 6, 1-10.

[14] Szyperski, T., Güntert, P., Otting, G. and Wüthrich, K. (1992) J. Mag. Res. 99, 552-560.

[15] Guntert, P., Mumenthaler, C. and Wuthrich, K. (1997) J. Mol. Biol. 273, 283-298.

[16] Koradi, R., Billeter, M. and Wuthrich, K. (1996) J. Mol. Graph. 14, 51-55.

[17] Hess, B., Kutzner, C., van der Spoel, D. and Lindahl, E. (2008) J. Chem. Theory Comp. 4, 435-447.

[18] Wishart, D.S., Bigam, C.G., Holm, A., Hodges, R.S. and Sykes, B.D. (1995) J. Biomol. NMR 5, 67-81.

[19] Mulder, F.A.A. and Filatov, M. (2010) Chem. Soc. Revs. 39, 578-590.

[20] Schwarzinger, S., Kroon, G.J.A., Foss, T.R., Chung, J., Wright, P.E. and Dyson, H.J. (2001) J. Am. Chem. Soc. 123, 2970-2978.

[21] Gorres, K.L. and Raines, R.T. (2010) Crit. Revs. Biochem. Mol. Biol. 45, 106-124.

[22] Macarthur, M.W. and Thornton, J.M. (1991) J. Mol. Biol. 218, 397-412.

[23] Adzhubei, A.A. and Sternberg, M.J.E. (1993) J. Mol. Biol. 229, 472-493.

[24] Shen, Y., Delaglio, F., Cornilescu, G. and Bax, A. (2009) J. Biomol. NMR 44, 213-223.

[25] Wüthrich, K. (1986) NMR of Proteins and Nucleic Acids, John Wiley & Sons, New York.

[26] Ösapay, K. and Case, D.A. (1994) J. Biomol. NMR 4, 215-230.

Figures

Figure-1. The sequence of SP. The normal single letter amino acid code is used, except for O* = *trans*-L-4-hydroxyproline. The regions associated with specific biological responses are also shown (JH = juvenile hormone).

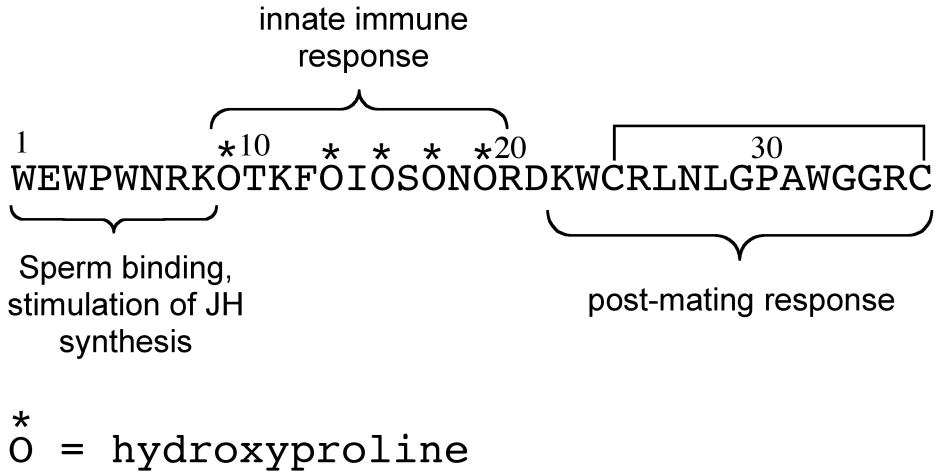


Figure-2. ^{1}H and ^{13}C Chemical shift deviations (CSDs) from random coil values for residues 1-36 in SP ($\Delta\delta=\delta^{observed} - \delta^{random}$) for $^{1}H\alpha$ (**A**), $^{13}C\alpha$ (**B**) and $^{13}C\beta$ (**C**) resonances in water solution. The asterik shows the positions of hydroxyproline (Hyp) in the sequence. Values for Gly are not included.

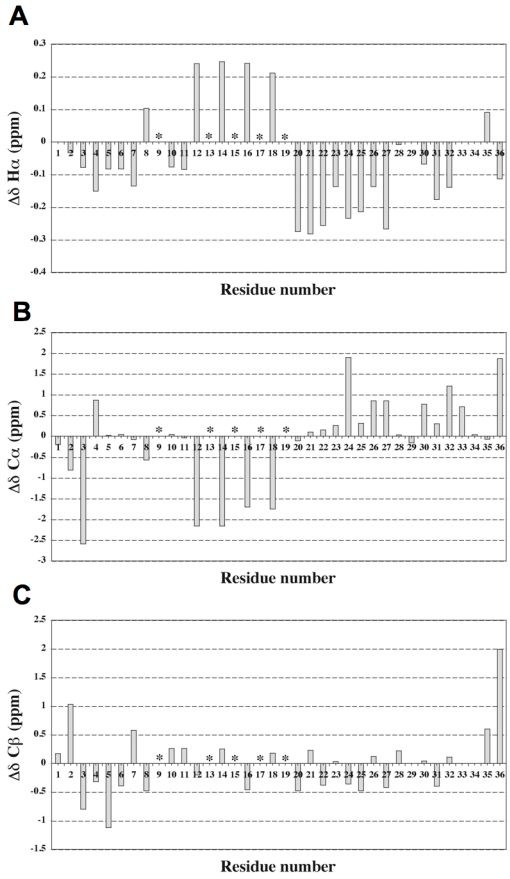


Figure-3. NOEs and $^{3}J_{HN\alpha}$ values observed for SP in (**A**) water and (**B**) water:TFE (1:1) solution. The letter "O" indicates hydroxyproline. NOEs shown to peptide NH groups are to the $CH_2(\delta)$ for P and O residues.

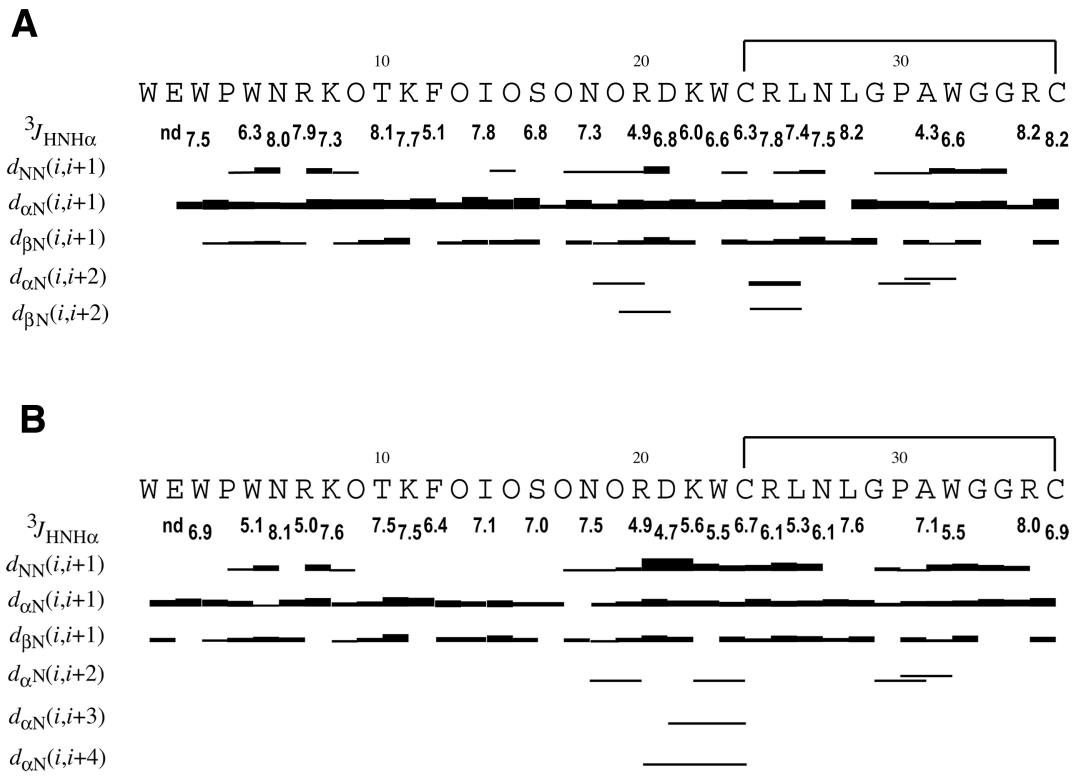


Figure-4. (**A**) Backbone representation and superposition of the 20 DYANA NMR structures of SP using the backbone heavy atoms of Phe^{12} to Asn^{18} (in blue), observed in TFE/water (*left*) and water (*right*) (the N-terminal region between Trp^{1} and Lys^{11} is shown in red, the C-terminal region between Hyp^{19} and Cys^{36} in green). (**B**) The Trp-rich N-terminal residues 1-8 in TFE/water superimposed (using the backbone heavy atoms) between Trp^{3} and Asn^{6} (in orange), showing the type I ß-turn at WPWN. (**C**) Representative side chain orientation in the turn region between 3 and 6. (**D**) Backbone representation and superposition of SP in TFE/water between Hyp^{19} and Cys^{36} (disulfide bond in yellow). (**E**) Ribbon representation of one solution structure of the region Hyp^{19} - Cys^{36} showing the nascent helical element. The side chains of Trp^{23}, Leu^{26}, Leu^{28} and Trp^{32} are arranged in a hydrophobic patch.

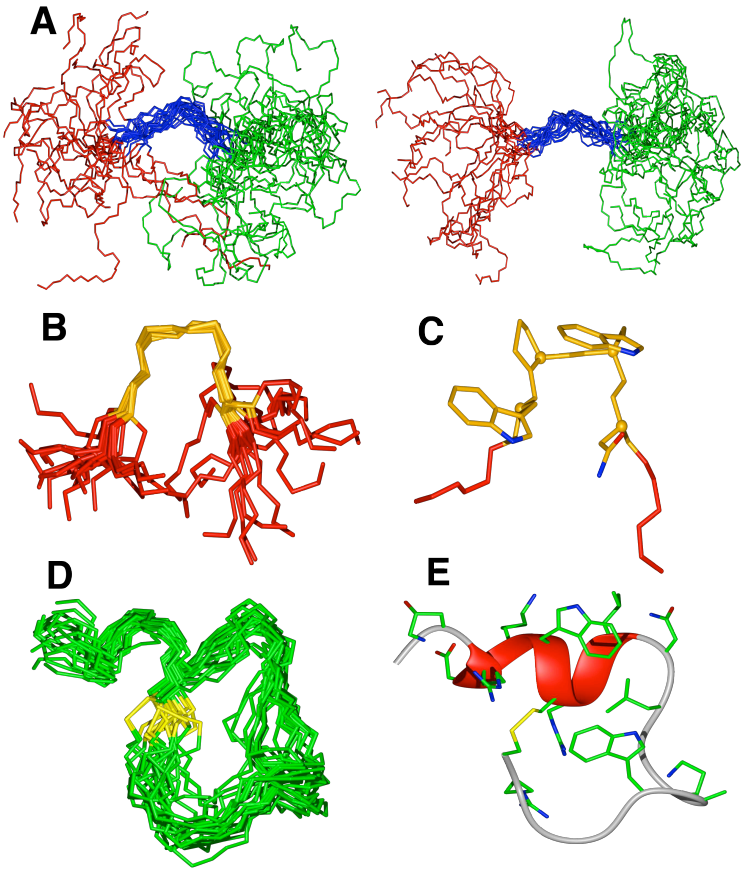


Table-1. Experimental distance restraints and statistics for the final 20 NMR structures calculated for SP in water and TFE/water.

	water	***TFE/water***
Number of NOE upper-distance limits	263	360
Intraresidue	101	127
Sequential	138	182
Medium- and long-range	24	51
Residual target function value ($Å^2$)	2.43 ± 0.40	1.49 ± 0.11
Residual NOE violations		
Number > 0.2 Å	32	15
Maximum (Å)	0.5	0.44
Mean rmsd values (Å) *between Glu^2 and Lys^8*		
backbone atoms	2.05 ± 0.46	1.62 ± 0.54
All heavy atoms	4.62 ± 0.79	3.55 ± 0.97
between Phe^{12} and Asn^{18}		
backbone atoms	1.96 ± 0.57	2.10 ± 0.59
All heavy atoms	3.51 ± 0.81	3.62 ± 0.84
between Lys^{22} and Cys^{36}		
backbone atoms	2.90 ± 1.31	1.97 ± 0.57
All heavy atoms	4.44 ± 1.61	3.10 ± 0.78

Supplementary data

NMR Studies of the Solution Conformation of the Sex Peptide from *Drosophila melanogaster*

Kerstin Moehle[a], Annabelle Freund[a], Eric Kubli[b] and John A. Robinson[a]*

[a] Chemistry Department, University of Zurich, Winterthurerstrasse 190, 8057 Zurich, Switzerland,

[b] Institute of Molecular Life Sciences, University of Zurich, Winterthurerstrasse 190, 8057 Zurich, Switzerland.

The following data are provided below:

1) Chemical shift assignments for SP in water (**Table S1**)

2) Chemical shift assignments for SP in TFE/water (**Table S2**)

3) A list of NOE-derived distance restraints in water (**Table S3**)

4) A list of NOE-derived distance restraints in TFE/water (**Table S4**)

5) ^{15}N-Chemical shift deviations from random coil (**Figure S1**)

6) MD simulation of SP with time-averaged NOE-derived distance restraints: ϕ/ψ backbone torsion angles (**Figure-S2**).

Table S1 ^{1}H , ^{13}C and ^{15}N chemical shifts (ppm) of SP in water (pH 5.0, 300 K).

1	Trp	CA	57.300
1	Trp	HA	4.168
1	Trp	CB	29.770
1	Trp	HB2	3.121
1	Trp	HB3	3.198
1	Trp	NE1	129.811
1	Trp	HD1	6.921
1	Trp	HE3	7.380
1	Trp	HE1	9.672
1	Trp	HZ3	6.730
1	Trp	HZ2	7.358
1	Trp	HH2	7.015
2	Glu	N	123.180
2	Glu	HN	8.088
2	Glu	CA	55.790
2	Glu	HA	4.259
2	Glu	CB	30.930
2	Glu	HB2	1.655
2	Glu	HB3	1.817
2	Glu	HG2	2.036
2	Glu	HG3	2.086
3	Trp	N	123.710
3	Trp	HN	8.143
3	Trp	CA	54.910
3	Trp	HA	4.614
3	Trp	CB	28.800
3	Trp	QB	2.573
3	Trp	NE1	129.169
3	Trp	HD1	7.187
3	Trp	HE3	7.643
3	Trp	HE1	10.176
3	Trp	HZ3	7.250
3	Trp	HZ2	7.507
3	Trp	HH2	7.260
4	Pro	CA	64.170
4	Pro	HA	4.290
4	Pro	CB	31.780
4	Pro	HB2	1.842
4	Pro	HB3	2.207
4	Pro	HG2	1.911
4	Pro	HG3	1.956
4	Pro	HD2	3.390
4	Pro	HD3	3.801
5	Trp	N	117.357
5	Trp	HN	7.371
5	Trp	CA	57.520
5	Trp	HA	4.624
5	Trp	CB	28.480
5	Trp	QB	3.309
5	Trp	NE1	130.395

```
 5 Trp  HD1    7.179
 5 Trp  HE3    7.563
 5 Trp  HE1   10.137
 5 Trp  HZ3    7.169
 5 Trp  HZ2    7.469
 5 Trp  HH2    7.206
 6 Asn    N  119.590
 6 Asn   HN    7.959
 6 Asn   CA   53.140
 6 Asn   HA    4.671
 6 Asn   CB   38.510
 6 Asn   QB    2.590
 6 Asn  ND2  111.560
 6 Asn HD21    7.570
 6 Asn HD22    6.875
 7 Arg    N  120.540
 7 Arg   HN    7.806
 7 Arg   CA   55.920
 7 Arg   HA    4.246
 7 Arg   CB   31.480
 7 Arg  HB2    1.658
 7 Arg  HB3    1.806
 7 Arg  HG2    1.515
 7 Arg  HG3    1.565
 7 Arg   QD    3.073
 8 Lys    N  123.810
 8 Lys   HN    8.291
 8 Lys   CA   55.630
 8 Lys   HA    4.477
 8 Lys   CB   32.620
 8 Lys  HB2    1.665
 8 Lys  HB3    1.765
 8 Lys   QG    1.401
 8 Lys   QD    1.657
 8 Lys   QE    2.978
 9 Hyp   CA   61.960
 9 Hyp   HA    4.582
 9 Hyp   CB   39.920
 9 Hyp  HB2    1.964
 9 Hyp  HB3    2.319
 9 Hyp   CG   72.660
 9 Hyp  HG2    4.555
 9 Hyp  HD2    3.689
 9 Hyp  HD3    3.766
10 Thr    N  118.840
10 Thr   HN    8.359
10 Thr   CA   61.840
10 Thr   HA    4.273
10 Thr   CB   70.060
10 Thr   HB    4.124
10 Thr  QG2    1.137
10 Thr  CG2   21.680
```

```
11 Lys    N  123.510
11 Lys   HN    8.171
11 Lys   CA   56.160
11 Lys   HA    4.276
11 Lys   CB   33.360
11 Lys  HB2    1.616
11 Lys  HB3    1.690
11 Lys  HG2    1.243
11 Lys  HG3    1.304
11 Lys   QD    1.626
11 Lys   QE    2.946
12 Phe    N  122.070
12 Phe   HN    8.226
12 Phe   CA   55.540
12 Phe   HA    4.902
12 Phe   CB   39.400
12 Phe  HB2    2.893
12 Phe  HB3    3.141
12 Phe   QD    7.294
12 Phe   QE    7.352
13 Hyp   CA   61.960
13 Hyp   HA    4.573
13 Hyp   CB   39.870
13 Hyp  HB2    2.014
13 Hyp  HB3    2.296
13 Hyp   CG   72.660
13 Hyp  HG2    4.571
13 Hyp  HD2    3.700
13 Hyp  HD3    3.820
14 Ile    N  123.240
14 Ile   HN    8.399
14 Ile   CA   58.940
14 Ile   HA    4.470
14 Ile   CB   39.050
14 Ile   HB    1.880
14 Ile  QG2    0.993
14 Ile  CG1   17.300
14 Ile HG12    1.229
14 Ile HG13    1.575
14 Ile  QD1    0.911
14 Ile  CD1   13.100
15 Hyp   CA   61.960
15 Hyp   HA    4.583
15 Hyp   CB   39.920
15 Hyp  HB2    2.040
15 Hyp  HB3    2.341
15 Hyp   CG   72.660
15 Hyp  HG2    4.585
15 Hyp  HD2    3.823
15 Hyp  HD3    3.936
16 Ser    N  118.680
16 Ser   HN    8.583
```

```
16 Ser   CA   56.600
16 Ser   HA    4.729
16 Ser   CB   63.340
16 Ser   QB    3.843
17 Hyp   CA   61.920
17 Hyp   HA    4.570
17 Hyp   CB   39.920
17 Hyp  HB2    2.023
17 Hyp  HB3    2.316
17 Hyp   CG   72.660
17 Hyp  HG2    4.570
17 Hyp   QD    3.831
18 Asn    N  121.030
18 Asn   HN    8.762
18 Asn   CA   51.350
18 Asn   HA    4.962
18 Asn   CB   39.080
18 Asn  HB2    2.762
18 Asn  HB3    2.896
18 Asn  ND2  113.270
18 Asn HD21    7.700
18 Asn HD22    6.976
19 Hyp   CA   63.060
19 Hyp   HA    4.515
19 Hyp   CB   40.070
19 Hyp  HB2    2.095
19 Hyp  HB3    2.396
19 Hyp  HG2    4.612
19 Hyp   QD    3.911
20 Arg    N  118.840
20 Arg   HN    8.356
20 Arg   CA   55.890
20 Arg   HA    4.111
20 Arg   CB   30.420
20 Arg  HB2    1.725
20 Arg  HB3    1.819
20 Arg   QG    1.617
21 Asp    N  119.900
21 Asp   HN    8.047
21 Asp   CA   54.300
21 Asp   HA    4.497
21 Asp   CB   41.330
21 Asp   QB    2.675
22 Lys    N  120.130
22 Lys   HN    7.992
22 Lys   CA   56.350
22 Lys   HA    4.103
22 Lys   CB   32.720
22 Lys   QB    1.700
22 Lys   QG    1.238
22 Lys   QD    1.575
22 Lys   QE    2.864
```

```
23 Trp    N  120.070
23 Trp   HN    7.995
23 Trp   CA   57.760
23 Trp   HA    4.554
23 Trp   CB   29.630
23 Trp  HB2    3.260
23 Trp  HB3    3.334
23 Trp  NE1  129.405
23 Trp  HD1    7.240
23 Trp  HE3    7.554
23 Trp  HE1   10.149
23 Trp  HZ3    7.113
24 Cys    N  118.950
24 Cys   HN    8.268
24 Cys   CA   57.300
24 Cys   HA    4.460
24 Cys   CB   40.740
24 Cys   QB    3.039
25 Arg    N  121.070
25 Arg   HN    8.165
25 Arg   CA   56.310
25 Arg   HA    4.164
25 Arg   CB   30.420
25 Arg  HB2    1.701
25 Arg  HB3    1.799
25 Arg   CG   29.250
25 Arg   QG    1.551
25 Arg   QD    3.101
26 Leu    N  119.870
26 Leu   HN    7.833
26 Leu   CA   55.950
26 Leu   HA    4.243
26 Leu   CB   42.520
26 Leu  HB2    1.522
26 Leu  HB3    1.650
26 Leu   HG    1.584
26 Leu  QD1    0.844
26 Leu  QD2    0.877
27 Asn    N  117.025
27 Asn   HN    8.078
27 Asn   CA   53.950
27 Asn   HA    4.481
27 Asn   CB   38.480
27 Asn  HB2    2.447
27 Asn  HB3    2.652
27 Asn  ND2  112.700
27 Asn HD21    7.248
27 Asn HD22    6.720
28 Leu    N  120.260
28 Leu   HN    7.976
28 Leu   CA   55.130
28 Leu   HA    4.362
```

```
28 Leu   CB   42.620
28 Leu  HB2    1.470
28 Leu  HB3    1.574
28 Leu   HG    1.471
28 Leu  QQD    0.725
29 Gly    N  108.780
29 Gly   HN    8.050
29 Gly   CA   44.943
29 Gly  HA1    3.976
29 Gly  HA2    4.245
30 Pro   CA   64.070
30 Pro   HA    4.369
30 Pro   CB   32.140
30 Pro  HB2    1.813
30 Pro  HB3    2.214
30 Pro   QG    1.962
30 Pro  HD2    3.586
30 Pro  HD3    3.685
31 Ala    N  122.230
31 Ala   HN    8.308
31 Ala   CA   52.800
31 Ala   HA    4.174
31 Ala   QB    1.248
31 Ala   CB   18.700
32 Trp    N  119.100
32 Trp   HN    7.720
32 Trp   CA   57.270
32 Trp   HA    4.562
32 Trp   CB   29.710
32 Trp  HB2    3.213
32 Trp  HB3    3.337
32 Trp  NE1  129.567
32 Trp  HD1    7.224
32 Trp  HE3    7.574
32 Trp  HE1   10.115
32 Trp  HZ3    7.116
32 Trp  HZ2    7.470
33 Gly    N  110.630
33 Gly   HN    8.105
33 Gly   CA   45.809
33 Gly  HA1    3.753
33 Gly  HA2    3.840
34 Gly    N  107.760
34 Gly   HN    7.382
34 Gly   CA   45.137
34 Gly   QA    3.740
35 Arg    N  119.120
35 Arg   HN    7.964
35 Arg   CA   55.930
35 Arg   HA    4.470
35 Arg   CB   31.500
35 Arg  HB2    1.763
```

```
35 Arg  HB3    1.899
35 Arg   QG    1.588
35 Arg   QD    3.125
36 Cys    N  125.123
36 Cys   HN    8.212
36 Cys   CA   58.710
36 Cys   HA    4.564
36 Cys   CB   43.090
36 Cys  HB2    3.040
36 Cys  HB3    3.254
```

Table S2. ^{1}H chemical shifts of the major *trans* form of SP in TFE/water (1:1) (pH 5.0, 300 K).

Residue	NH	H-C(α)	H-C(β)	Others
Trp[1]	-	4.24	3.19, 3.29	H(ε^1) 8.92; H(δ^1) 6.85; H(ε^3) 7.51; H(ζ^3) 6.90; H(η^2) 7.15; H(ζ^2) 7.40
Glu[2]	7.82	4.44	1.86,1.91	CH_2(γ) 2.04, 2.04
Trp[3]	8.14	4.66	2.05, 2.34	H(ε^1) 10.04; H(δ^1) 7.36; H(ε^3) 7.65; H(ζ^3) 7.42; H(η^2) 7.41; H(ζ^2) 7.63
Pro[4]	-	4.31	2.05, 2.42	CH_2(γ) 2.14, 2.14; CH_2(δ) 3.41, 3.91
Trp[5]	6.77	4.83	3.45, 3.54	H(ε^1) 9.93; H(δ^1) 7.22; H(ε^3) 7.65; H(ζ^3) 7.30; H(η^2) 7.09; H(ζ^2) 7.00
Asn[6]	7.89	5.05	2.84, 3.03	NH(ε) 7.73; NH(ζ) 7.01
Arg[7]	7.82	4.45	1.89, 2.00	CH_2(γ) 1.71, 1.77; CH_2(δ) 3.27, 3.27; NH(ε) 7.18; NH_2(η) -, -
Lys[8]	8.29	4.73	1.88, 1.98	CH_2(γ) 1.62, 1.62; CH_2(δ) 1.87, 1.87; CH_2(ε) 3.16, 3.16; NH_3^+(ζ) -
Hyp[9]	-	4.67	2.17, 2.51	CH(γ) 4.55; CH_2(δ) 3.89, 4.00
Thr[10]	8.32	4.48	4.30	CH_3(γ) 1.33
Lys[11]	8.16	4.46	1.79, 1.88	CH_2(γ) 1.45, 1.49; CH_2(δ) 1.79, 1.79; CH_2(ε) 3.11, 3.11; NH_3^+(ζ) -
Phe[12]	8.13	5.05	3.09, 3.26	H(δ) 7.42, 7.42; H(ε) 7.47, 7.47; H(ζ) -
Hyp[13]	-	4.72	2.20, 2.42	CH(γ) 4.67; CH_2(δ) 3.69, 3.95
Ile[14]	8.11	4.66	2.02	$CH_2(\gamma^1)$ 1.36, 1.74; $CH_3(\gamma^2)$ 1.16; $CH_3(\delta^1)$ 1.06
Hyp[15]	-	4.75	2.22, 2.50	CH(γ) -; CH_2(δ) 3.94, 4.11
Ser[16]	8.50	4.95	4.03, 4.03	-
Hyp[17]	-		2.18, 2.50	CH(γ) 4.69; CH_2(δ) 4.01, 4.01
Asn[18]	8.81	5.15	3.02, 3.15	NH(ε) 7.86; NH(ζ) 6.98
Hyp[19]	-	4.66	2.21, 2.60	CH(γ) 4.61; CH_2(δ) 4.10, 4.10
Arg[20]	8.32	4.29	1.97, 2.09	CH_2(γ) 1.87, 1.87; CH_2(δ) 3.34, 3.34; NH(ε) 7.38; NH_2(η) -, -
Asp[21]	8.11	4.66	2.96, 3.03	-
Lys[22]	7.91	4.16	1.78, 1.97	CH_2(γ) 1.46, 1.55; CH_2(δ) 1.79, 1.79; CH_2(ε) 3.03, 3.03; NH_3^+(ζ) -
Trp[23]	8.17	4.44	3.46, 3.61	H(ε^1) 9.99; H(δ^1) 7.31; H(ε^3) 7.69; H(ζ^3) 7.22; H(η^2) 7.31; H(ζ^2) 7.55
Cys[24]	8.62	4.38	3.36, 3.36	
Arg[25]	8.18	4.16	1.88, 1.94	CH_2(γ) 1.65, 1.74; CH_2(δ) 3.16, 3.16; NH(ε) 7.38; NH_2(η) -, -
Leu[26]	8.06	4.29	1.60, 1.80	CH(γ) 1.81; $CH_3(\delta^1)$ 0.98; $CH_3(\delta^2)$ 0.98
Asn[27]	7.82	4.60	2.28, 2.51	NH(ε) 6.79; NH(ζ) 5.84
Leu[28]	8.09	4.52	1.53, 1.63	CH(γ) 1.47; $CH_3(\delta^1)$ 0.85; $CH_3(\delta^2)$ 0.85
Gly[29]	7.93	4.02, 4.42	-	-
Pro[30]	-	4.51	2.02, 2.39	CH_2(γ) 2.13, 2.13; CH_2(δ) 3.73, 3.92
Ala[31]	8.28	4.45	1.48	

Trp^{32}	7.80	4.64	3.39, 3.56	$H(\varepsilon^1)$ 10.00; $H(\delta^1)$ 7.40; $H(\varepsilon^3)$ 7.72; $H(\zeta^3)$ 7.26; $H(\eta^2)$ 7.34; $H(\zeta^2)$ 7.58
Gly^{33}	8.11	3.79, 3.96	-	-
Gly^{34}	7.43	3.80, 3.95	-	-
Arg^{35}	7.90	4.64	1.96, 2.07	$CH_2(\gamma)$ 1.78, 1.78; $CH_2(\delta)$ 3.27, 3.27; $NH(\varepsilon)$ 7.22; $NH_2(\eta)$ -, -
Cys^{36}	8.17		3.29, 3.46	-

Table S3. NOE upper distance limits in water.

INTRARESIDUAL NOE UPPER-DISTANCE LIMITS

1	TRP	HA	1	TRP	HD1	5.50
1	TRP	HA	1	TRP	HE3	5.50
2	GLU-	HA	2	GLU-	HB2	2.55
2	GLU-	HA	2	GLU-	HB3	2.55
3	TRP	HN	3	TRP	HD1	5.22
3	TRP	HA	3	TRP	HD1	5.50
3	TRP	HA	3	TRP	HE3	4.38
5	TRP	HA	5	TRP	HD1	4.48
5	TRP	HA	5	TRP	HE3	4.88
7	ARG+	HN	7	ARG+	HB2	3.67
7	ARG+	HN	7	ARG+	HB3	3.67
7	ARG+	HN	7	ARG+	QB	3.36
7	ARG+	HN	7	ARG+	HG2	4.97
7	ARG+	HN	7	ARG+	HG3	4.97
7	ARG+	HN	7	ARG+	QG	4.23
7	ARG+	HA	7	ARG+	HG2	3.30
7	ARG+	HA	7	ARG+	HG3	3.30
7	ARG+	HA	7	ARG+	QD	4.67
8	LYS+	HN	8	LYS+	HB2	3.24
8	LYS+	HN	8	LYS+	HB3	3.24
8	LYS+	HN	8	LYS+	QG	6.38
8	LYS+	HA	8	LYS+	HB2	2.90
8	LYS+	HA	8	LYS+	HB3	2.90
8	LYS+	HA	8	LYS+	QB	2.68
10	THR	HN	10	THR	HB	3.42
11	LYS+	HN	11	LYS+	HA	2.90
11	LYS+	HN	11	LYS+	HB2	3.30
11	LYS+	HN	11	LYS+	HB3	3.30
11	LYS+	HN	11	LYS+	QB	2.89
11	LYS+	HN	11	LYS+	HG2	5.50
11	LYS+	HN	11	LYS+	HG3	5.50
11	LYS+	HA	11	LYS+	QG	3.82
12	PHE	HN	12	PHE	HB2	3.73
12	PHE	HN	12	PHE	HB3	3.73
12	PHE	HN	12	PHE	QB	3.38
14	ILE	HN	14	ILE	HA	2.90
14	ILE	HN	14	ILE	HB	3.14
14	ILE	HN	14	ILE	HG12	4.45
14	ILE	HN	14	ILE	HG13	4.45
14	ILE	HN	14	ILE	QD1	6.53
14	ILE	HA	14	ILE	HB	2.90
18	ASN	HN	18	ASN	HB2	3.67
18	ASN	HN	18	ASN	HB3	3.67
18	ASN	HN	18	ASN	QB	3.50
18	ASN	HA	18	ASN	HB2	3.05
18	ASN	HA	18	ASN	HB3	3.05
20	ARG+	HN	20	ARG+	HB2	3.36
20	ARG+	HN	20	ARG+	HB3	3.36

```
20 ARG+ HN     20 ARG+ QB      3.18
20 ARG+ HN     20 ARG+ QG      5.73
20 ARG+ HA     20 ARG+ QD      6.38
21 ASP- HN     21 ASP- HA      2.77
22 LYS+ HN     22 LYS+ QG      5.76
23 TRP  HN     23 TRP  HB2     3.11
23 TRP  HN     23 TRP  HB3     3.11
23 TRP  HN     23 TRP  QB      2.92
23 TRP  HN     23 TRP  HD1     4.88
23 TRP  HN     23 TRP  HE3     5.50
23 TRP  HA     23 TRP  HD1     4.29
23 TRP  HA     23 TRP  HE3     4.45
23 TRP  HB2    23 TRP  HE3     3.79
23 TRP  HB3    23 TRP  HE3     3.79
23 TRP  QB     23 TRP  HE3     3.63
25 ARG+ HN     25 ARG+ HA      2.74
25 ARG+ HN     25 ARG+ HB2     2.99
25 ARG+ HN     25 ARG+ HB3     2.90
25 ARG+ HN     25 ARG+ QG      5.54
25 ARG+ HA     25 ARG+ HB2     2.65
26 LEU  HN     26 LEU  HB2     3.67
26 LEU  HN     26 LEU  HB3     3.67
26 LEU  HN     26 LEU  QB      3.42
26 LEU  HN     26 LEU  HG      4.35
26 LEU  HN     26 LEU  QD1     6.53
26 LEU  HN     26 LEU  QD2     6.53
26 LEU  HA     26 LEU  QD1     5.50
26 LEU  HA     26 LEU  QD2     5.50
27 ASN  HN     27 ASN  HA      2.86
27 ASN  HN     27 ASN  HB2     3.45
27 ASN  HN     27 ASN  HB3     3.45
28 LEU  HN     28 LEU  HB2     3.14
28 LEU  HN     28 LEU  HB3     2.83
28 LEU  HN     28 LEU  QQD     7.63
28 LEU  HA     28 LEU  HB2     2.93
28 LEU  HA     28 LEU  HB3     2.96
29 GLY  HN     29 GLY  HA2     2.93
29 GLY  HN     29 GLY  HA1     2.93
32 TRP  HN     32 TRP  HB2     3.42
32 TRP  HN     32 TRP  HB3     3.42
32 TRP  HA     32 TRP  HD1     4.32
32 TRP  HA     32 TRP  HE3     4.01
32 TRP  HB2    32 TRP  HE3     3.95
32 TRP  HB3    32 TRP  HE3     3.95
35 ARG+ HN     35 ARG+ HA      2.86
35 ARG+ HN     35 ARG+ HB2     3.83
35 ARG+ HN     35 ARG+ HB3     3.83
35 ARG+ HN     35 ARG+ QB      3.38
35 ARG+ HN     35 ARG+ QG      5.26
35 ARG+ HA     35 ARG+ QD      4.27
36 CYSS HN     36 CYSS HB3     3.30
36 CYSS HN     36 CYSS HB2     3.30
```

```
 36 CYSS HN     36 CYSS QB      3.10
SUM = 101


 SEQUENTIAL NOE UPPER-DISTANCE LIMITS

  2 GLU- HA      3 TRP  HN      3.14
  3 TRP  HA      4 PRO  HD2     3.05
  3 TRP  HA      4 PRO  HD3     3.05
  3 TRP  HA      4 PRO  QD      2.59
  3 TRP  QB      4 PRO  HD2     6.31
  3 TRP  QB      4 PRO  HD3     6.31
  3 TRP  QB      4 PRO  QD      6.06
  3 TRP  HE3     4 PRO  HD2     5.19
  3 TRP  HE3     4 PRO  HD3     5.19
  3 TRP  HE3     4 PRO  QD      4.76
  4 PRO  HA      5 TRP  HN      3.14
  4 PRO  QB      5 TRP  HN      4.48
  4 PRO  QD      5 TRP  HN      5.69
  5 TRP  HN      6 ASN  HN      3.42
  5 TRP  HA      6 ASN  HN      3.30
  5 TRP  QB      6 ASN  HN      4.52
  5 TRP  HD1     6 ASN  HN      5.31
  5 TRP  HE3     6 ASN  HN      5.50
  6 ASN  HA      7 ARG+ HN      3.36
  7 ARG+ HN      8 LYS+ HN      3.64
  7 ARG+ HA      8 LYS+ HN      2.74
  8 LYS+ HN      9 HYPR QD      6.20
  8 LYS+ HA      9 HYPR HD2     3.17
  8 LYS+ HA      9 HYPR HD3     3.17
  8 LYS+ HA      9 HYPR QD      2.83
  8 LYS+ HB2     9 HYPR HD2     5.50
  8 LYS+ HB2     9 HYPR HD3     5.50
  8 LYS+ HB3     9 HYPR HD2     5.50
  8 LYS+ HB3     9 HYPR HD3     5.50
  8 LYS+ QG      9 HYPR HD2     6.38
  8 LYS+ QG      9 HYPR HD3     6.38
  9 HYPR HA     10 THR  HN      2.77
  9 HYPR HB2    10 THR  HN      4.04
  9 HYPR HB3    10 THR  HN      4.04
 10 THR  HA     11 LYS+ HN      2.86
 10 THR  HB     11 LYS+ HN      3.30
 10 THR  QG2    11 LYS+ HN      6.53
 11 LYS+ HA     12 PHE  HN      2.65
 11 LYS+ HA     12 PHE  QD      7.62
 12 PHE  HA     13 HYPR HD2     3.45
 12 PHE  HA     13 HYPR HD3     3.45
 12 PHE  HA     13 HYPR QD      3.07
 12 PHE  HB2    13 HYPR HG2     5.50
 12 PHE  HB2    13 HYPR HD2     5.38
 12 PHE  HB2    13 HYPR HD3     5.38
 12 PHE  HB3    13 HYPR HG2     5.50
```

```
12 PHE  HB3    13 HYPR HD2     5.38
12 PHE  HB3    13 HYPR HD3     5.38
12 PHE  QB     13 HYPR QD      4.39
12 PHE  QD     13 HYPR HD2     7.62
12 PHE  QD     13 HYPR HD3     7.62
13 HYPR HA     14 ILE  HN      2.49
13 HYPR HB2    14 ILE  HN      4.17
13 HYPR HB3    14 ILE  HN      4.17
13 HYPR QB     14 ILE  HN      3.92
14 ILE  HN     15 HYPR HD2     5.50
14 ILE  HN     15 HYPR HD3     5.50
14 ILE  HA     15 HYPR HD2     2.71
14 ILE  HA     15 HYPR HD3     2.71
14 ILE  HB     15 HYPR HD2     4.88
14 ILE  HB     15 HYPR HD3     4.88
14 ILE  HB     15 HYPR QD      4.12
14 ILE  QG2    15 HYPR HD2     5.66
14 ILE  QG2    15 HYPR HD3     5.66
14 ILE  QG2    15 HYPR QD      5.34
15 HYPR HA     16 SER  HN      2.59
15 HYPR HB2    16 SER  HN      3.89
15 HYPR HB3    16 SER  HN      3.89
15 HYPR QB     16 SER  HN      3.72
16 SER  HA     17 HYPR QD      4.69
17 HYPR HA     18 ASN  HN      2.86
17 HYPR HB2    18 ASN  HN      4.29
17 HYPR HB3    18 ASN  HN      4.29
17 HYPR QB     18 ASN  HN      4.07
18 ASN  HN     19 HYPR QD      6.39
18 ASN  HA     19 HYPR QD      3.82
18 ASN  HB2    19 HYPR QD      5.87
18 ASN  HB3    19 HYPR QD      5.87
19 HYPR HA     20 ARG+ HN      2.80
19 HYPR HB2    20 ARG+ HN      4.26
19 HYPR HB3    20 ARG+ HN      4.26
19 HYPR QB     20 ARG+ HN      4.01
20 ARG+ HN     21 ASP- HN      3.14
20 ARG+ HA     21 ASP- HN      2.99
21 ASP- HA     22 LYS+ HN      2.77
21 ASP- QB     22 LYS+ HN      4.11
22 LYS+ HA     23 TRP  HN      3.17
22 LYS+ HA     23 TRP  HD1     5.50
23 TRP  HN     24 CYSS HN      5.50
23 TRP  HA     24 CYSS HN      2.83
23 TRP  HB2    24 CYSS HN      3.89
23 TRP  HB3    24 CYSS HN      3.89
23 TRP  QB     24 CYSS HN      3.71
23 TRP  HD1    24 CYSS HN      5.50
23 TRP  HD1    24 CYSS HA      4.63
23 TRP  HE3    24 CYSS HN      5.50
23 TRP  HE3    24 CYSS HA      4.57
23 TRP  HE3    24 CYSS QB      6.38
```

```
 24 CYSS HA     25 ARG+ HN      2.80
 24 CYSS QB     25 ARG+ HN      4.36
 25 ARG+ HN     26 LEU  HN      5.50
 25 ARG+ HA     26 LEU  HN      3.39
 25 ARG+ HB2    26 LEU  HN      3.79
 25 ARG+ HB3    26 LEU  HN      3.45
 26 LEU  HN     27 ASN  HN      4.35
 26 LEU  HA     27 ASN  HN      2.90
 26 LEU  HB2    27 ASN  HN      3.24
 26 LEU  HB3    27 ASN  HN      3.24
 26 LEU  QB     27 ASN  HN      2.96
 26 LEU  QQD    27 ASN  HN      8.10
 28 LEU  HA     29 GLY  HN      2.77
 28 LEU  HB2    29 GLY  HN      3.55
 28 LEU  HB3    29 GLY  HN      3.11
 28 LEU  QQD    29 GLY  HN      7.63
 29 GLY  HN     30 PRO  QD      5.74
 29 GLY  HA2    30 PRO  HD2     3.73
 29 GLY  HA2    30 PRO  HD3     3.73
 29 GLY  HA1    30 PRO  HD2     3.73
 29 GLY  HA1    30 PRO  HD3     3.73
 29 GLY  QA     30 PRO  QD      3.00
 30 PRO  HA     31 ALA  HN      3.02
 30 PRO  QB     31 ALA  HN      4.70
 30 PRO  HD2    31 ALA  HN      5.50
 30 PRO  HD3    31 ALA  HN      5.50
 30 PRO  QD     31 ALA  HN      5.35
 31 ALA  HN     32 TRP  HN      3.64
 31 ALA  HA     32 TRP  HN      3.36
 31 ALA  QB     32 TRP  HD1     8.00
 32 TRP  HN     33 GLY  HN      3.98
 32 TRP  HA     33 GLY  HN      3.05
 32 TRP  HB2    33 GLY  HN      4.11
 32 TRP  HB3    33 GLY  HN      4.11
 32 TRP  QB     33 GLY  HN      3.89
 33 GLY  HN     34 GLY  HN      3.76
 33 GLY  HA2    34 GLY  HN      3.14
 33 GLY  HA1    34 GLY  HN      3.14
 33 GLY  QA     34 GLY  HN      2.84
 35 ARG+ HA     36 CYSS HN      2.68
SUM = 138


 MEDIUM and LONG RANGE NOE UPPER-DISTANCE LIMITS

  3 TRP  QB      5 TRP  HN      5.32
  3 TRP  QB      5 TRP  HE1     6.38
 10 THR  QG2    12 PHE  HN      6.53
 10 THR  QG2    12 PHE  QD      8.65
 18 ASN  HN     21 ASP- QB      5.38
 18 ASN  HA     20 ARG+ HN      5.50
 19 HYPR QB     21 ASP- HN      5.70
```

```
 20 ARG+ HA     23 TRP  HD1     5.50
 20 ARG+ HA     23 TRP  HE3     5.50
 20 ARG+ HA     23 TRP  HE1     5.50
 23 TRP  HB2    28 LEU  QQD     7.63
 23 TRP  HB3    28 LEU  QQD     7.63
 23 TRP  HE3    28 LEU  QQD     7.63
 24 CYSS HA     26 LEU  HN      3.30
 24 CYSS QB     26 LEU  HN      6.31
 24 CYSS QB     28 LEU  QQD     8.51
 24 CYSS QB     32 TRP  HE3     6.38
 28 LEU  HB2    32 TRP  HD1     4.63
 28 LEU  HB3    32 TRP  HD1     4.38
 28 LEU  QQD    32 TRP  HB2     7.63
 28 LEU  QQD    32 TRP  HB3     7.63
 28 LEU  QQD    32 TRP  HD1     7.63
 28 LEU  QQD    32 TRP  HE3     7.63
 28 LEU  QQD    32 TRP  HE1     7.63
 28 LEU  QQD    32 TRP  HZ3     7.63
 29 GLY  QA     31 ALA  HN      4.55
 30 PRO  HA     32 TRP  HN      4.14
SUM = 27
```

Table S4. NOE upper distance limits in TFE/water.

```
INTRARESIDUAL NOE UPPER-DISTANCE LIMITS

  1 TRP  HA      1 TRP  HD1     6.00
  1 TRP  HA      1 TRP  HE3     5.39
  1 TRP  HB2     1 TRP  HD1     3.70
  1 TRP  HB3     1 TRP  HD1     3.70
  1 TRP  QB      1 TRP  HD1     3.54
  2 GLU- HN      2 GLU- HB2     3.55
  2 GLU- HN      2 GLU- HB3     3.55
  2 GLU- HN      2 GLU- QG      6.70
  2 GLU- HA      2 GLU- HB2     2.76
  2 GLU- HA      2 GLU- HB3     2.76
  2 GLU- HA      2 GLU- QB      2.49
  3 TRP  HN      3 TRP  QB      3.69
  3 TRP  HN      3 TRP  HD1     4.78
  3 TRP  HA      3 TRP  HD1     6.00
  3 TRP  HA      3 TRP  HE3     4.09
  5 TRP  HN      5 TRP  HB2     4.02
  5 TRP  HN      5 TRP  HB3     4.02
  5 TRP  HN      5 TRP  QB      3.72
  5 TRP  HN      5 TRP  HE3     4.52
  5 TRP  HA      5 TRP  HD1     6.00
  5 TRP  HA      5 TRP  HE3     6.00
  6 ASN  HN      6 ASN  HB2     3.62
  6 ASN  HN      6 ASN  HB3     3.62
  6 ASN  HN      6 ASN  QB      3.42
```

6	ASN	HA	6	ASN	HD21	6.00
6	ASN	HA	6	ASN	HD22	6.00
7	ARG+	HN	7	ARG+	HB2	3.52
7	ARG+	HN	7	ARG+	HB3	3.52
7	ARG+	HN	7	ARG+	HG2	5.60
7	ARG+	HN	7	ARG+	HG3	5.60
7	ARG+	HN	7	ARG+	QG	5.00
7	ARG+	HA	7	ARG+	HB2	2.94
7	ARG+	HA	7	ARG+	HB3	2.94
7	ARG+	HA	7	ARG+	HG2	3.55
7	ARG+	HA	7	ARG+	HG3	3.55
7	ARG+	HA	7	ARG+	QG	3.26
7	ARG+	HB2	7	ARG+	HE	6.00
7	ARG+	HB3	7	ARG+	HE	6.00
8	LYS+	HN	8	LYS+	HB2	3.16
8	LYS+	HN	8	LYS+	HB3	3.16
8	LYS+	HN	8	LYS+	QB	2.92
8	LYS+	HN	8	LYS+	QG	6.26
10	THR	HN	10	THR	HB	3.16
11	LYS+	HN	11	LYS+	HA	2.76
11	LYS+	HN	11	LYS+	HB2	3.12
11	LYS+	HN	11	LYS+	HB3	3.12
11	LYS+	HN	11	LYS+	QB	2.88
11	LYS+	HN	11	LYS+	HG2	4.70
11	LYS+	HN	11	LYS+	HG3	4.70
11	LYS+	HN	11	LYS+	QG	4.06
11	LYS+	HA	11	LYS+	QG	3.71
12	PHE	HN	12	PHE	HB2	3.37
12	PHE	HN	12	PHE	HB3	3.37
12	PHE	HN	12	PHE	QB	3.02
14	ILE	HN	14	ILE	HB	3.00
14	ILE	HN	14	ILE	HG12	4.45
14	ILE	HN	14	ILE	HG13	4.45
14	ILE	HN	14	ILE	QG1	3.76
14	ILE	HN	14	ILE	QD1	6.74
18	ASN	HN	18	ASN	HB2	3.59
18	ASN	HN	18	ASN	HB3	3.59
18	ASN	HN	18	ASN	QB	3.40
20	ARG+	HN	20	ARG+	HB2	3.30
20	ARG+	HN	20	ARG+	HB3	3.30
20	ARG+	HN	20	ARG+	QB	3.07
20	ARG+	HN	20	ARG+	QG	5.65
21	ASP-	HN	21	ASP-	HB2	3.48
21	ASP-	HN	21	ASP-	HB3	3.48
21	ASP-	HN	21	ASP-	QB	3.19
22	LYS+	HN	22	LYS+	HB2	3.23
22	LYS+	HN	22	LYS+	HB3	3.23
22	LYS+	HN	22	LYS+	QB	2.96
22	LYS+	HN	22	LYS+	HG2	4.88
22	LYS+	HN	22	LYS+	HG3	4.88
22	LYS+	HN	22	LYS+	QG	4.64
22	LYS+	HA	22	LYS+	HG2	3.84

```
 22 LYS+ HA     22 LYS+ HG3     3.84
 22 LYS+ HA     22 LYS+ QG      3.68
 22 LYS+ QB     22 LYS+ QE      6.64
 23 TRP  HN     23 TRP  HB2     3.41
 23 TRP  HN     23 TRP  HB3     3.41
 23 TRP  HN     23 TRP  QB      3.13
 23 TRP  HN     23 TRP  HD1     4.67
 23 TRP  HN     23 TRP  HE3     5.00
 23 TRP  HA     23 TRP  HD1     4.56
 23 TRP  HA     23 TRP  HE3     4.13
 23 TRP  HB2    23 TRP  HE3     4.02
 23 TRP  HB3    23 TRP  HE3     4.02
 23 TRP  QB     23 TRP  HE1     5.18
 25 ARG+ HN     25 ARG+ HB2     3.30
 25 ARG+ HN     25 ARG+ HB3     3.30
 25 ARG+ HA     25 ARG+ HB2     2.90
 25 ARG+ HA     25 ARG+ HB3     2.90
 25 ARG+ HA     25 ARG+ HG2     3.52
 25 ARG+ HA     25 ARG+ HG3     3.52
 25 ARG+ HA     25 ARG+ QG      3.12
 25 ARG+ HA     25 ARG+ HE      5.89
 25 ARG+ QG     25 ARG+ HE      3.85
 26 LEU  HN     26 LEU  HB2     2.98
 26 LEU  HN     26 LEU  HB3     3.23
 26 LEU  HN     26 LEU  QQD     7.56
 26 LEU  HA     26 LEU  HB2     2.65
 26 LEU  HA     26 LEU  HB3     2.98
 27 ASN  HN     27 ASN  HB2     3.59
 27 ASN  HN     27 ASN  HB3     3.59
 27 ASN  HN     27 ASN  QB      3.39
 27 ASN  HN     27 ASN  QD2     5.32
 28 LEU  HN     28 LEU  QQD     8.02
 31 ALA  HN     31 ALA  HA      2.83
 32 TRP  HN     32 TRP  HB2     3.19
 32 TRP  HN     32 TRP  HB3     3.19
 32 TRP  HN     32 TRP  HD1     4.60
 32 TRP  HA     32 TRP  HD1     5.10
 32 TRP  HA     32 TRP  HE3     4.45
 32 TRP  HB2    32 TRP  HE3     3.88
 32 TRP  HB3    32 TRP  HE3     3.88
 32 TRP  QB     32 TRP  HE3     3.69
 35 ARG+ HN     35 ARG+ HB2     3.41
 35 ARG+ HN     35 ARG+ HB3     3.41
 35 ARG+ HN     35 ARG+ QB      3.13
 35 ARG+ HN     35 ARG+ QG      4.68
 35 ARG+ HB2    35 ARG+ HE      5.53
 35 ARG+ HB3    35 ARG+ HE      5.53
 35 ARG+ QB     35 ARG+ HE      5.37
 36 CYSS HN     36 CYSS HB3     3.23
 36 CYSS HN     36 CYSS HB2     3.23
 36 CYSS HN     36 CYSS QB      3.00
SUM = 127
```

SEQUENTIAL NOE UPPER-DISTANCE LIMITS

1	TRP	HA	2	GLU-	HN	3.41
1	TRP	HB2	2	GLU-	HN	4.27
1	TRP	HB3	2	GLU-	HN	4.27
1	TRP	QB	2	GLU-	HN	4.07
1	TRP	HE3	2	GLU-	HA	6.00
2	GLU-	HA	3	TRP	HN	3.05
3	TRP	HA	4	PRO	HD2	3.30
3	TRP	HA	4	PRO	HD3	3.30
3	TRP	HA	4	PRO	QD	2.96
3	TRP	HB2	4	PRO	HD2	6.99
3	TRP	HB2	4	PRO	HD3	6.99
3	TRP	HB3	4	PRO	HD2	6.99
3	TRP	HB3	4	PRO	HD3	6.99
3	TRP	QB	4	PRO	QD	6.02
4	PRO	HA	5	TRP	HN	3.59
4	PRO	HB2	5	TRP	HN	4.67
4	PRO	HB3	5	TRP	HN	4.67
4	PRO	QB	5	TRP	HN	4.05
4	PRO	QB	5	TRP	HZ3	6.02
4	PRO	QG	5	TRP	HE3	6.71
4	PRO	QG	5	TRP	HZ3	6.89
4	PRO	HD2	5	TRP	HE3	4.38
4	PRO	HD3	5	TRP	HE3	4.38
4	PRO	QD	5	TRP	HN	5.05
4	PRO	QD	5	TRP	HE3	4.09
5	TRP	HN	6	ASN	HN	3.62
5	TRP	QB	6	ASN	HN	4.61
5	TRP	HD1	6	ASN	HN	6.00
5	TRP	HD1	6	ASN	HA	6.00
5	TRP	HE3	6	ASN	HN	4.67
6	ASN	HA	7	ARG+	HN	3.30
6	ASN	HB2	7	ARG+	HN	4.24
6	ASN	HB3	7	ARG+	HN	4.24
6	ASN	QB	7	ARG+	HN	3.88
7	ARG+	HN	8	LYS+	HN	3.41
7	ARG+	HA	8	LYS+	HN	2.98
8	LYS+	HN	9	HYPR	QD	6.13
8	LYS+	HA	9	HYPR	HD2	4.13
8	LYS+	HA	9	HYPR	HD3	4.13
8	LYS+	HA	9	HYPR	QD	3.80
8	LYS+	HB2	9	HYPR	HD2	6.00
8	LYS+	HB2	9	HYPR	HD3	6.00
8	LYS+	HB3	9	HYPR	HD2	6.00
8	LYS+	HB3	9	HYPR	HD3	6.00
8	LYS+	QB	9	HYPR	QD	4.97
8	LYS+	QG	9	HYPR	HD2	6.88
8	LYS+	QG	9	HYPR	HD3	6.88
9	HYPR	HB2	10	THR	HN	4.38

```
 9 HYPR HB3    10 THR  HN      4.38
 9 HYPR QB     10 THR  HN      4.16
10 THR  HA     11 LYS+ HN      2.83
10 THR  HB     11 LYS+ HN      3.08
10 THR  QG2    11 LYS+ HN      5.80
11 LYS+ HA     12 PHE  HN      2.94
12 PHE  HA     13 HYPR HD2     3.30
12 PHE  HA     13 HYPR HD3     3.30
12 PHE  HA     13 HYPR QD      2.96
12 PHE  HB2    13 HYPR HD2     5.14
12 PHE  HB2    13 HYPR HD3     5.14
12 PHE  HB3    13 HYPR HD2     5.14
12 PHE  HB3    13 HYPR HD3     5.14
12 PHE  QB     13 HYPR QD      4.25
12 PHE  QD     13 HYPR HD2     8.12
12 PHE  QD     13 HYPR HD3     8.12
12 PHE  QD     13 HYPR QD      7.66
13 HYPR HB2    14 ILE  HN      3.95
13 HYPR HB3    14 ILE  HN      3.95
13 HYPR QB     14 ILE  HN      3.76
14 ILE  HA     15 HYPR HD2     3.59
14 ILE  HA     15 HYPR HD3     3.59
14 ILE  HA     15 HYPR QD      3.08
14 ILE  HB     15 HYPR HD2     4.06
14 ILE  HB     15 HYPR HD3     4.06
14 ILE  HB     15 HYPR QD      3.23
14 ILE  QG2    15 HYPR HD2     5.91
14 ILE  QG2    15 HYPR HD3     5.91
14 ILE  QG2    15 HYPR QD      5.52
14 ILE  QG1    15 HYPR QD      7.77
14 ILE  QD1    15 HYPR HD2     7.03
14 ILE  QD1    15 HYPR HD3     7.03
15 HYPR HB2    16 SER  HN      4.34
15 HYPR HB3    16 SER  HN      4.34
15 HYPR QB     16 SER  HN      4.16
16 SER  HA     17 HYPR QD      3.98
17 HYPR HB2    18 ASN  HN      4.63
17 HYPR HB3    18 ASN  HN      4.63
17 HYPR QB     18 ASN  HN      4.42
18 ASN  HN     19 HYPR QD      6.89
18 ASN  HA     19 HYPR QD      4.16
18 ASN  HB2    19 HYPR QD      6.14
18 ASN  HB3    19 HYPR QD      6.14
18 ASN  QB     19 HYPR QD      5.57
19 HYPR HB2    20 ARG+ HN      4.38
19 HYPR HB3    20 ARG+ HN      4.38
19 HYPR QB     20 ARG+ HN      4.06
20 ARG+ HN     21 ASP- HN      2.40
20 ARG+ HA     21 ASP- HN      3.12
20 ARG+ QB     21 ASP- HN      4.22
21 ASP- HN     22 LYS+ HN      2.44
21 ASP- HA     22 LYS+ HN      3.48
```

21	ASP-	HB2	22	LYS+	HN	3.80
21	ASP-	HB3	22	LYS+	HN	3.80
21	ASP-	QB	22	LYS+	HN	3.63
22	LYS+	HN	23	TRP	HN	3.20
22	LYS+	HA	23	TRP	HN	3.50
22	LYS+	HB2	23	TRP	HB2	6.81
22	LYS+	HB2	23	TRP	HB3	6.81
22	LYS+	HB3	23	TRP	HB2	6.81
22	LYS+	HB3	23	TRP	HB3	6.81
22	LYS+	QB	23	TRP	QB	5.18
23	TRP	HN	24	CYSS	HN	3.59
23	TRP	HB2	24	CYSS	HN	3.70
23	TRP	HB3	24	CYSS	HN	3.70
23	TRP	QB	24	CYSS	HN	3.54
23	TRP	HD1	24	CYSS	HN	6.00
23	TRP	HE3	24	CYSS	HN	5.50
23	TRP	HE3	24	CYSS	HA	4.45
23	TRP	HE3	24	CYSS	QB	6.84
24	CYSS	HN	25	ARG+	HN	3.48
24	CYSS	HA	25	ARG+	HN	3.12
24	CYSS	QB	25	ARG+	HN	4.46
25	ARG+	HN	26	LEU	HN	3.00
25	ARG+	HA	26	LEU	HN	3.60
25	ARG+	HB2	26	LEU	HN	3.52
25	ARG+	HB3	26	LEU	HN	3.52
25	ARG+	QB	26	LEU	HN	3.26
26	LEU	HN	27	ASN	HN	3.20
26	LEU	HA	27	ASN	HN	3.50
26	LEU	HB2	27	ASN	HN	3.66
26	LEU	HB2	27	ASN	HD21	6.00
26	LEU	HB2	27	ASN	HD22	6.00
26	LEU	HB2	27	ASN	QD2	5.16
26	LEU	HB3	27	ASN	HN	3.44
26	LEU	HB3	27	ASN	HD21	6.00
26	LEU	HB3	27	ASN	HD22	6.00
26	LEU	QQD	27	ASN	HN	8.13
27	ASN	HN	28	LEU	QQD	8.13
27	ASN	HA	28	LEU	HN	3.19
27	ASN	HB2	28	LEU	HN	4.30
27	ASN	HB2	28	LEU	QQD	8.13
27	ASN	HB3	28	LEU	HN	4.30
27	ASN	HB3	28	LEU	QQD	8.13
28	LEU	HA	29	GLY	HN	3.37
28	LEU	HB2	29	GLY	HN	3.95
28	LEU	HB3	29	GLY	HN	3.95
28	LEU	QB	29	GLY	HN	3.50
28	LEU	QQD	29	GLY	HN	8.13
29	GLY	HN	30	PRO	HD2	4.45
29	GLY	HN	30	PRO	HD3	4.45
29	GLY	HA2	30	PRO	HD2	5.23
29	GLY	HA2	30	PRO	HD3	5.23
29	GLY	HA1	30	PRO	HD2	5.23

```
 29 GLY  HA1    30 PRO  HD3     5.23
 29 GLY  QA     30 PRO  QD      4.19
 30 PRO  HA     31 ALA  HN      3.70
 30 PRO  HB2    31 ALA  HN      4.42
 30 PRO  HB3    31 ALA  HN      4.42
 30 PRO  QB     31 ALA  HN      3.88
 30 PRO  HD2    31 ALA  HN      5.28
 30 PRO  HD3    31 ALA  HN      5.28
 30 PRO  QD     31 ALA  HN      4.96
 31 ALA  HN     32 TRP  HN      3.37
 31 ALA  HN     32 TRP  HD1     6.00
 31 ALA  HA     32 TRP  HN      3.62
 31 ALA  HA     32 TRP  HD1     5.80
 31 ALA  QB     32 TRP  HD1     6.60
 32 TRP  HN     33 GLY  HN      3.20
 32 TRP  HA     33 GLY  HN      3.50
 32 TRP  HB2    33 GLY  HN      3.41
 32 TRP  HB3    33 GLY  HN      3.41
 32 TRP  HD1    33 GLY  HN      6.00
 32 TRP  HE3    33 GLY  HA2     6.00
 32 TRP  HE3    33 GLY  HA1     6.00
 32 TRP  HE3    33 GLY  QA      5.56
 33 GLY  HN     34 GLY  HN      3.48
 33 GLY  HA2    34 GLY  HN      3.23
 33 GLY  HA1    34 GLY  HN      3.23
 33 GLY  QA     34 GLY  HN      3.03
 34 GLY  HN     35 ARG+ HN      3.84
 34 GLY  HA2    35 ARG+ HN      3.34
 34 GLY  HA1    35 ARG+ HN      3.34
 35 ARG+ HA     36 CYSS HN      2.98
SUM = 182


 MEDIUM and LONG RANGE NOE UPPER-DISTANCE LIMITS

  3 TRP  HN      6 ASN  QB      4.86
  3 TRP  HB2     5 TRP  HE1     6.00
  3 TRP  HB3     5 TRP  HE1     6.00
  3 TRP  QB      5 TRP  HN      5.22
  3 TRP  QB      5 TRP  HD1     6.88
  3 TRP  QB      5 TRP  HZ2     6.88
  3 TRP  QB      6 ASN  HN      5.01
  3 TRP  HD1     6 ASN  QD2     4.78
 10 THR  QG2    12 PHE  HN      6.96
 10 THR  QG2    12 PHE  QD      9.15
 10 THR  QG2    12 PHE  QE      9.15
 18 ASN  HA     20 ARG+ HN      4.60
 18 ASN  HB2    20 ARG+ HN      4.31
 18 ASN  HB3    20 ARG+ HN      4.31
 18 ASN  QB     20 ARG+ HN      4.07
 20 ARG+ HA     23 TRP  HD1     5.68
 20 ARG+ HA     23 TRP  HE3     4.92
```

```
 20 ARG+ HA     24 CYSS HN      4.38
 21 ASP- HA     24 CYSS HN      4.70
 22 LYS+ HA     24 CYSS HN      4.13
 23 TRP  HA     27 ASN  HD21    5.50
 23 TRP  HA     27 ASN  HD22    5.50
 23 TRP  HA     27 ASN  QD2     5.32
 23 TRP  HD1    26 LEU  QQD     8.13
 23 TRP  HD1    28 LEU  QQD     8.13
 23 TRP  HE3    26 LEU  QQD     8.13
 23 TRP  HE3    28 LEU  QQD     8.13
 23 TRP  HZ3    26 LEU  QQD     8.13
 23 TRP  HZ3    28 LEU  QQD     8.13
 23 TRP  HH2    28 LEU  QQD     8.13
 24 CYSS QB     26 LEU  HN      6.41
 24 CYSS QB     28 LEU  QQD     9.01
 26 LEU  HN     28 LEU  QQD     8.13
 26 LEU  QQD    28 LEU  HN      8.13
 28 LEU  HA     32 TRP  HD1     5.32
 28 LEU  QB     32 TRP  HD1     6.63
 28 LEU  QQD    32 TRP  HN      8.13
 28 LEU  QQD    32 TRP  HB2     8.13
 28 LEU  QQD    32 TRP  HB3     8.13
 28 LEU  QQD    32 TRP  QB      7.90
 28 LEU  QQD    32 TRP  HD1     8.13
 28 LEU  QQD    32 TRP  HE3     8.13
 28 LEU  QQD    32 TRP  HE1     8.13
 28 LEU  QQD    32 TRP  HZ2     8.13
 28 LEU  QQD    36 CYSS HN      8.13
 28 LEU  QQD    36 CYSS HB3     8.13
 28 LEU  QQD    36 CYSS HB2     8.13
 28 LEU  QQD    36 CYSS QB      7.99
 29 GLY  QA     31 ALA  HN      5.08
 30 PRO  HA     32 TRP  HN      3.70
 32 TRP  HE3    36 CYSS QB      6.48
SUM = 51
```

Figure-S1. ^{15}N-Chemical shift deviations from random coil for residues 1-36 in SP ($\Delta\delta=\delta^{observed} - \delta^{random}$) in water solution. The asterik shows the positions of hydroxy-proline (Hyp) in the sequence.

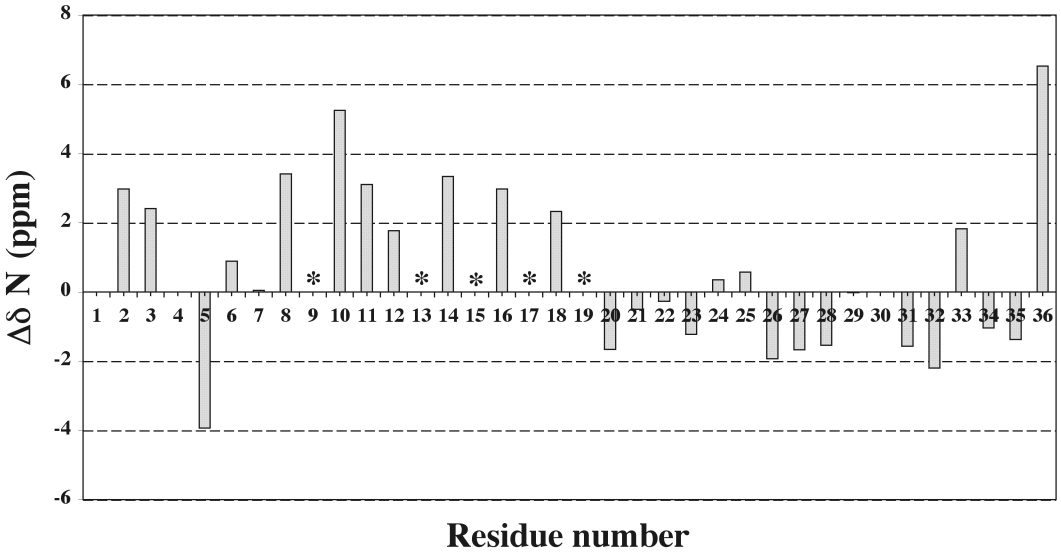
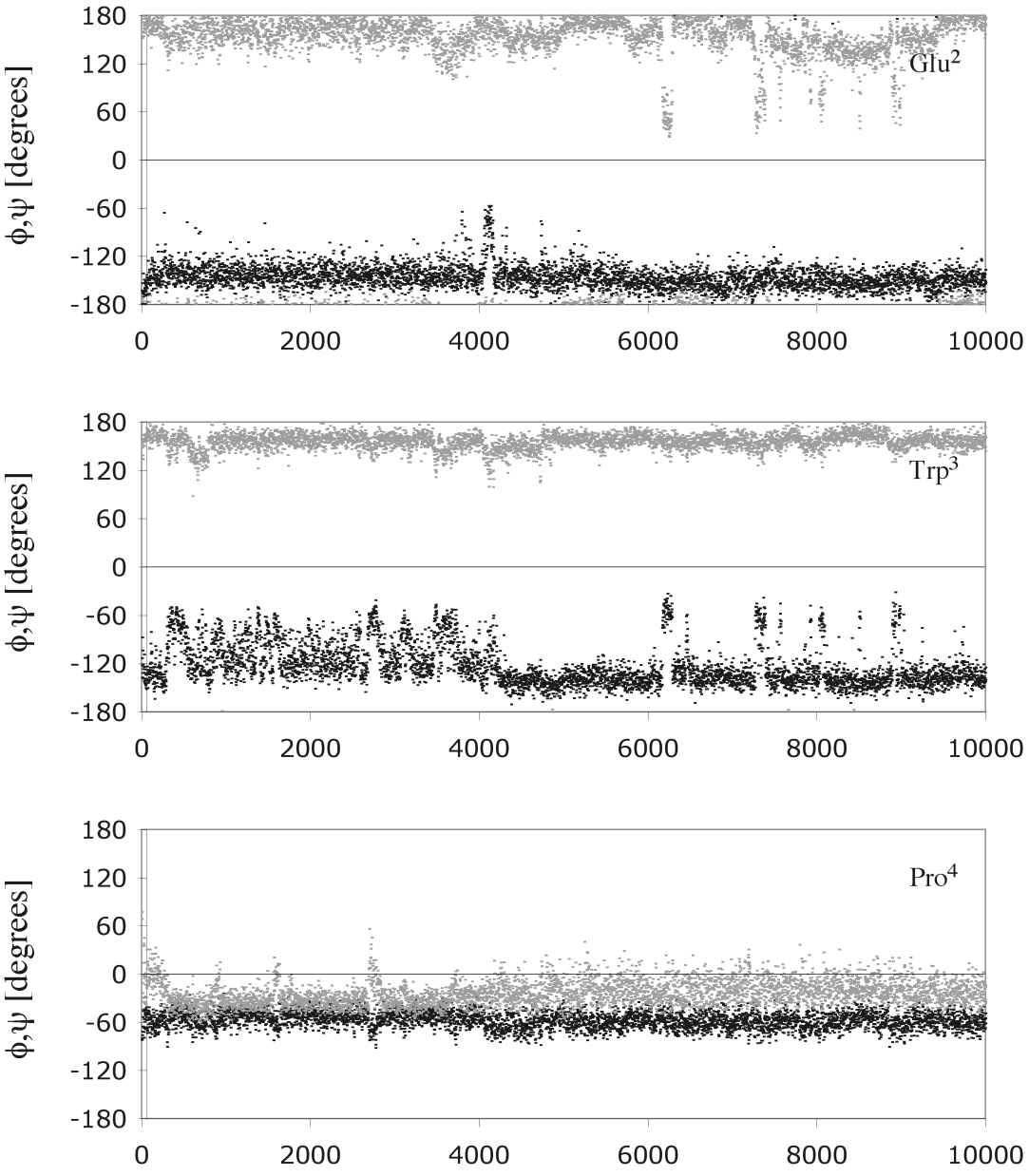
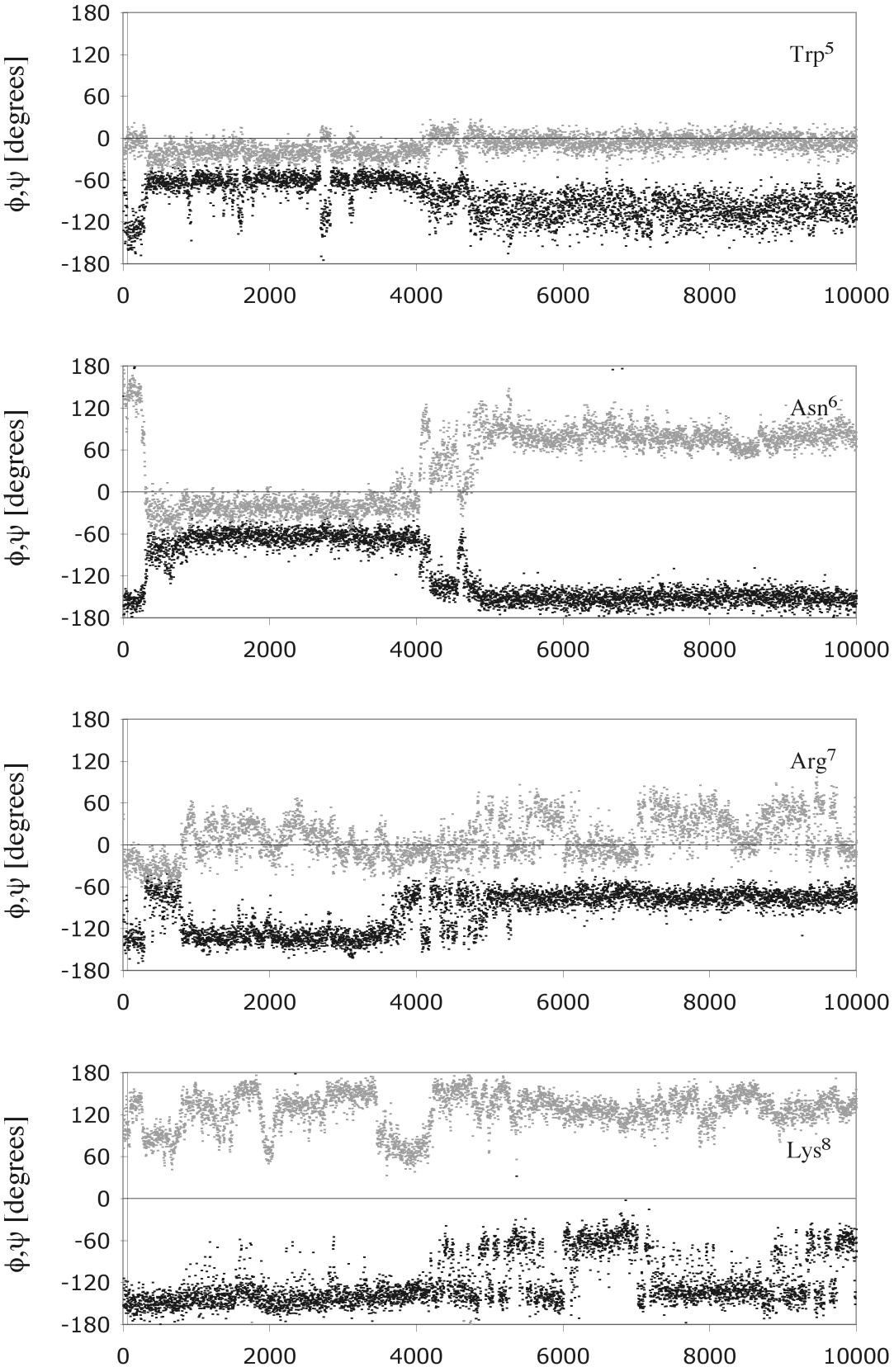
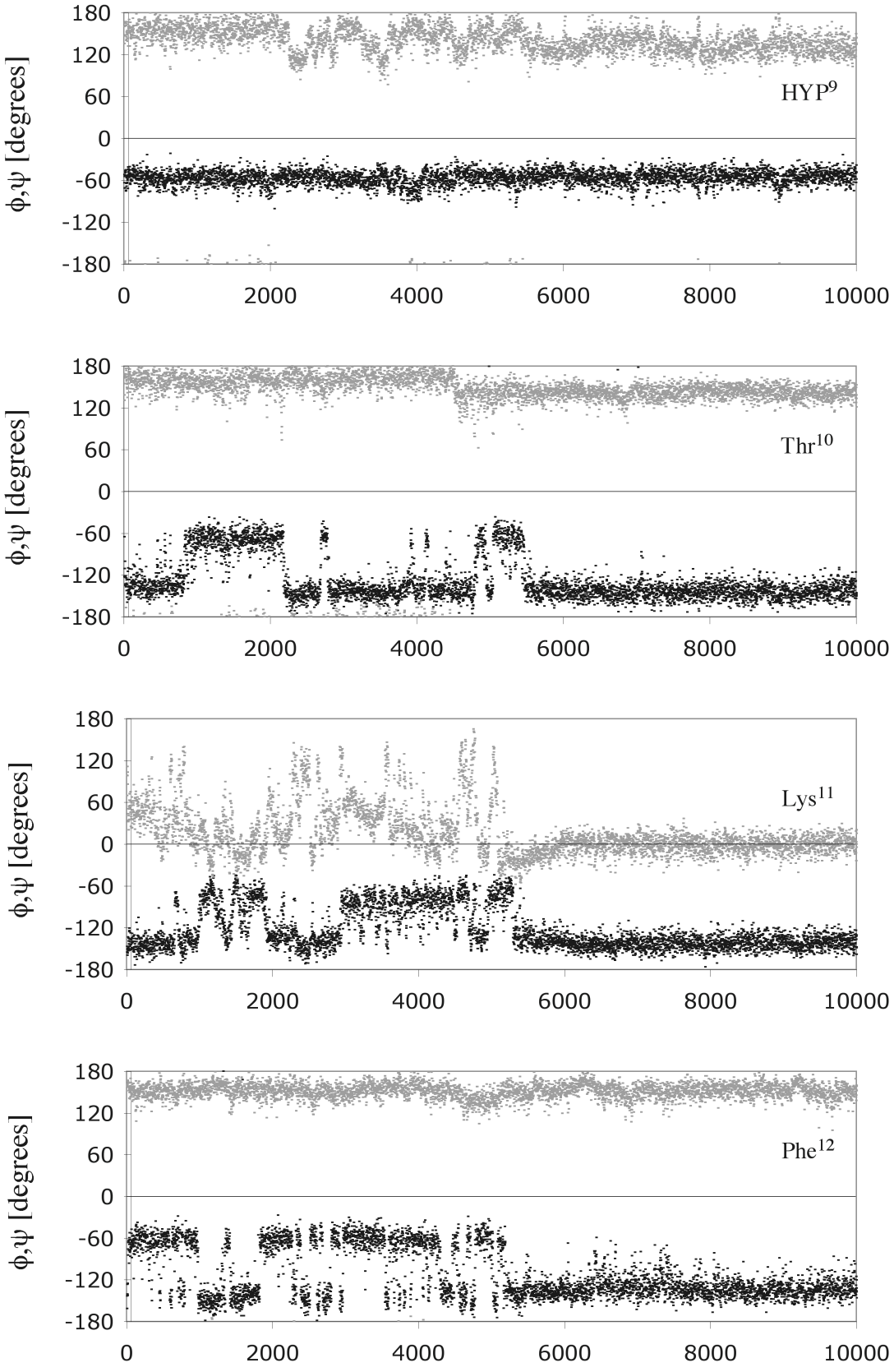
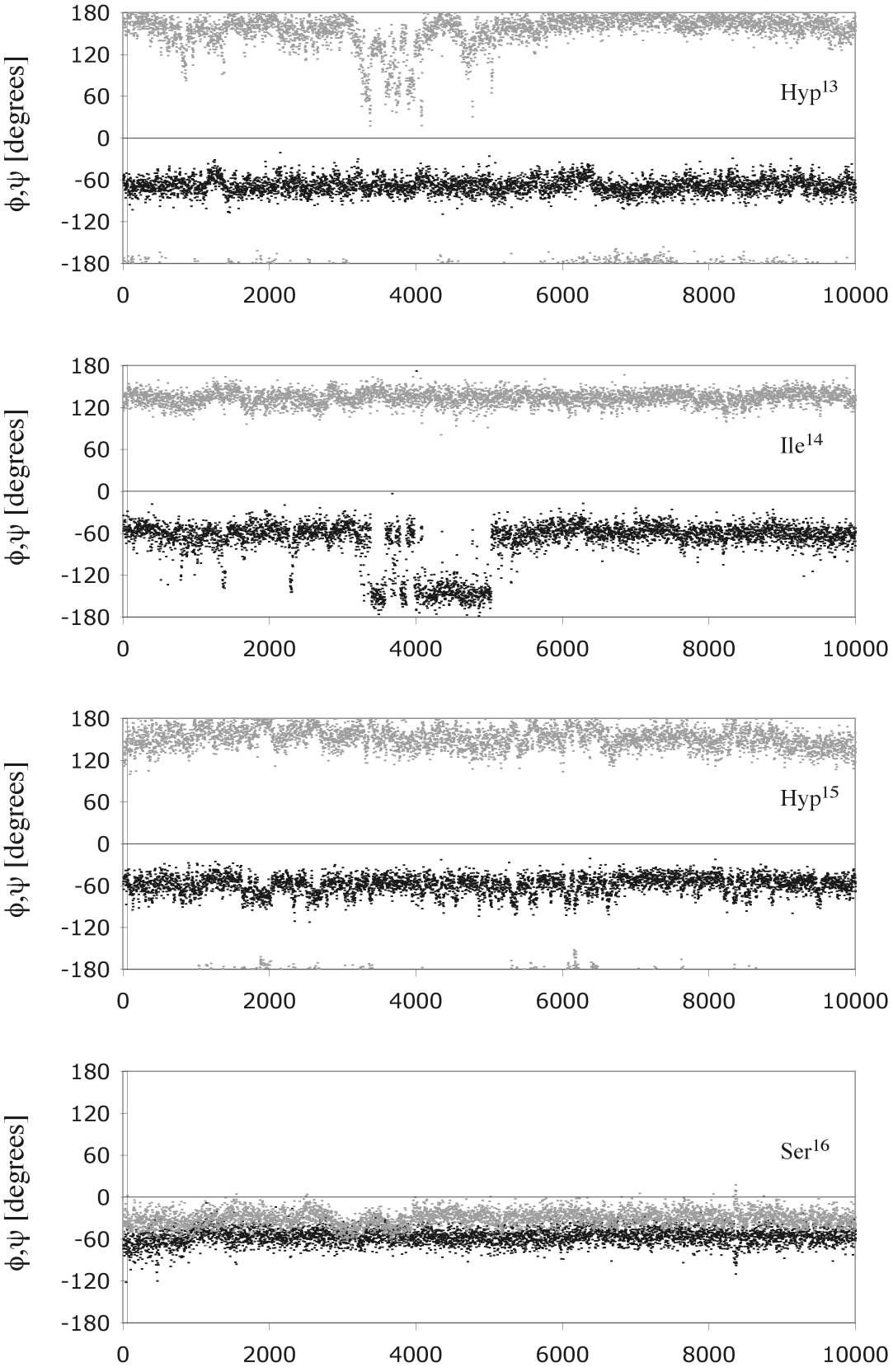


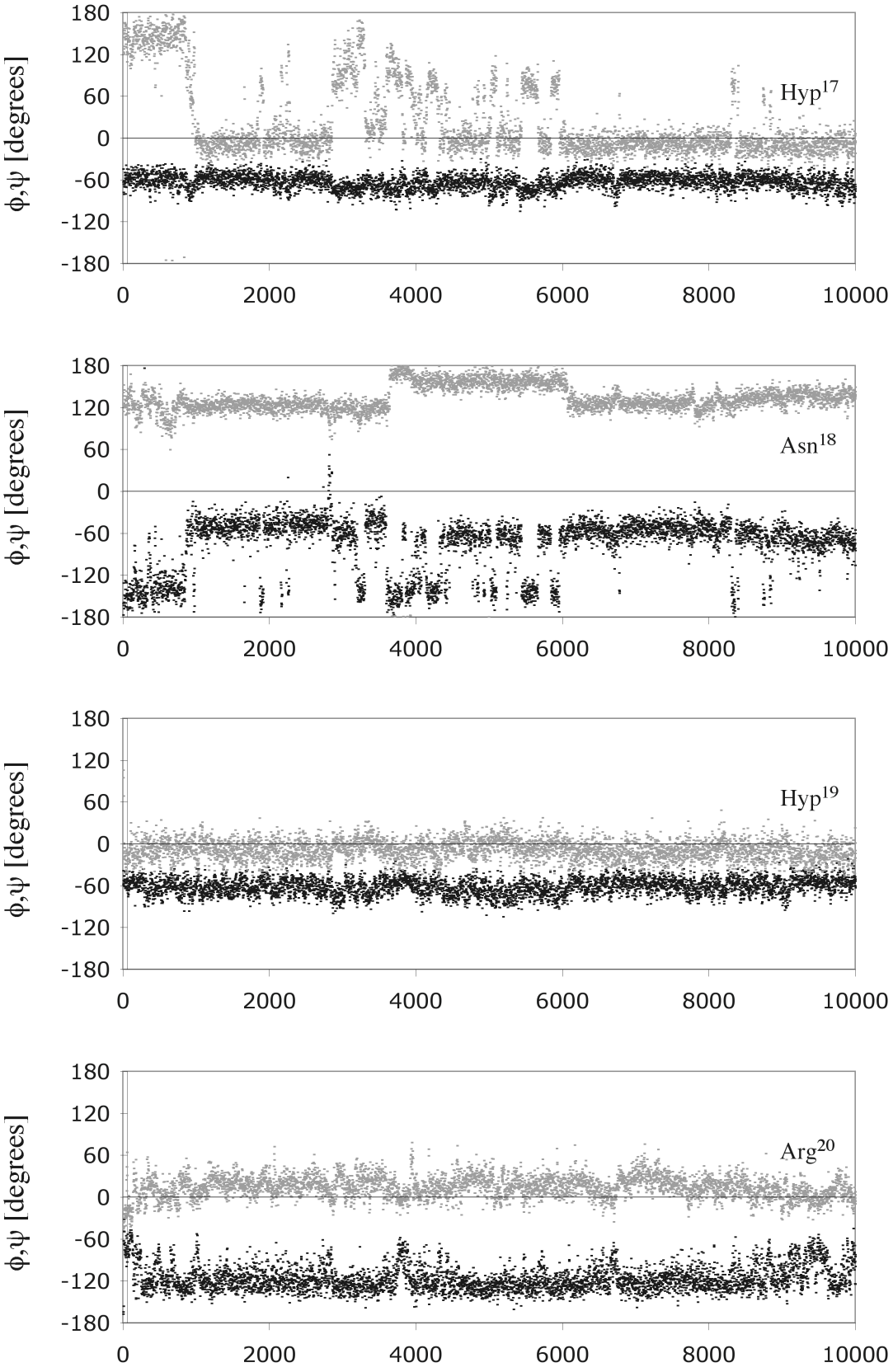
Figure-S2. Results from the MD simulation of SP with time-averaged NOE-derived distance restraints. The ϕ (black lines), ψ (gray lines) backbone torsion angles of each residue in SP, from Glu^2 to Arg^{35}, as a function of simulation time in ps.

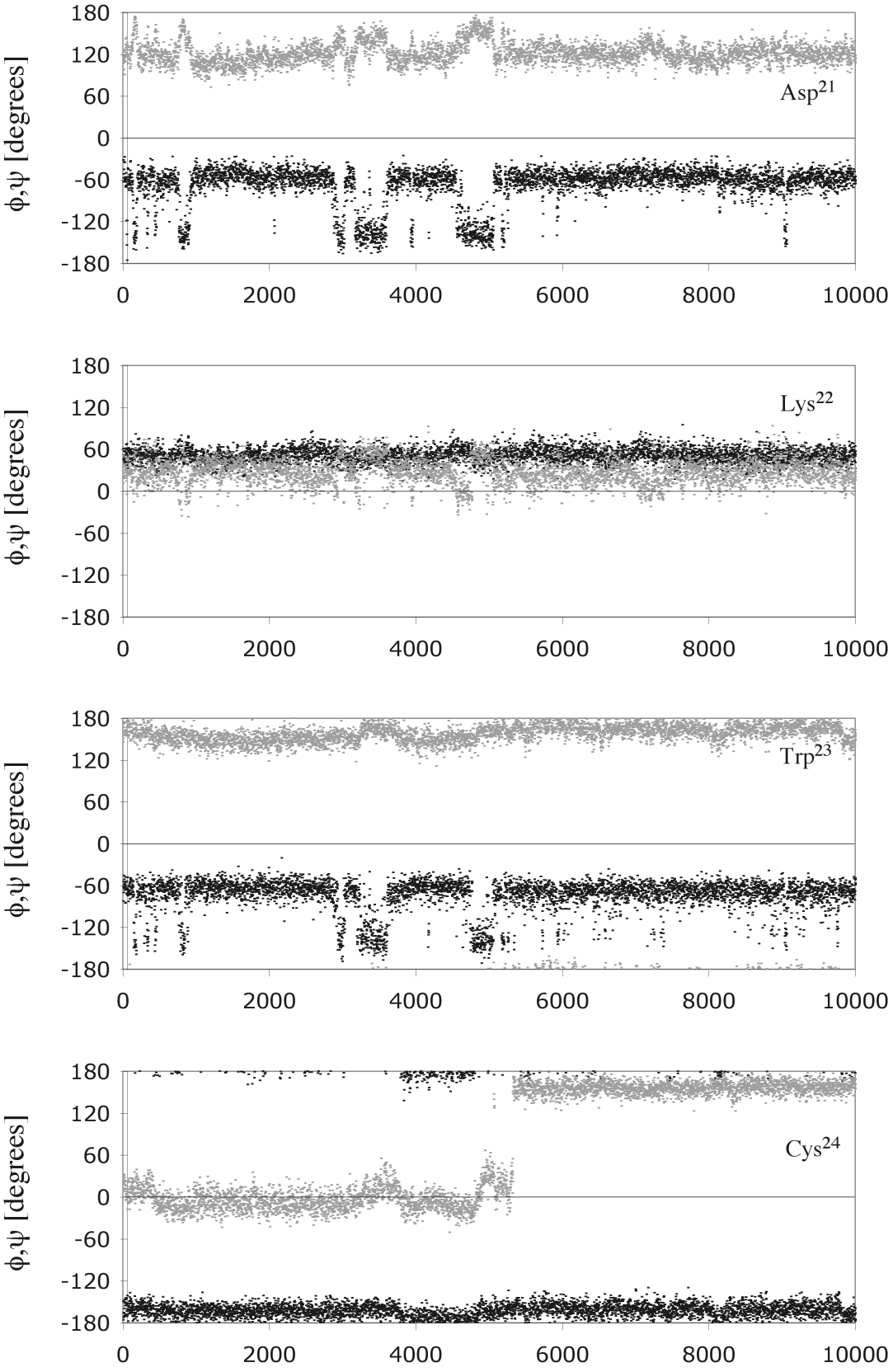


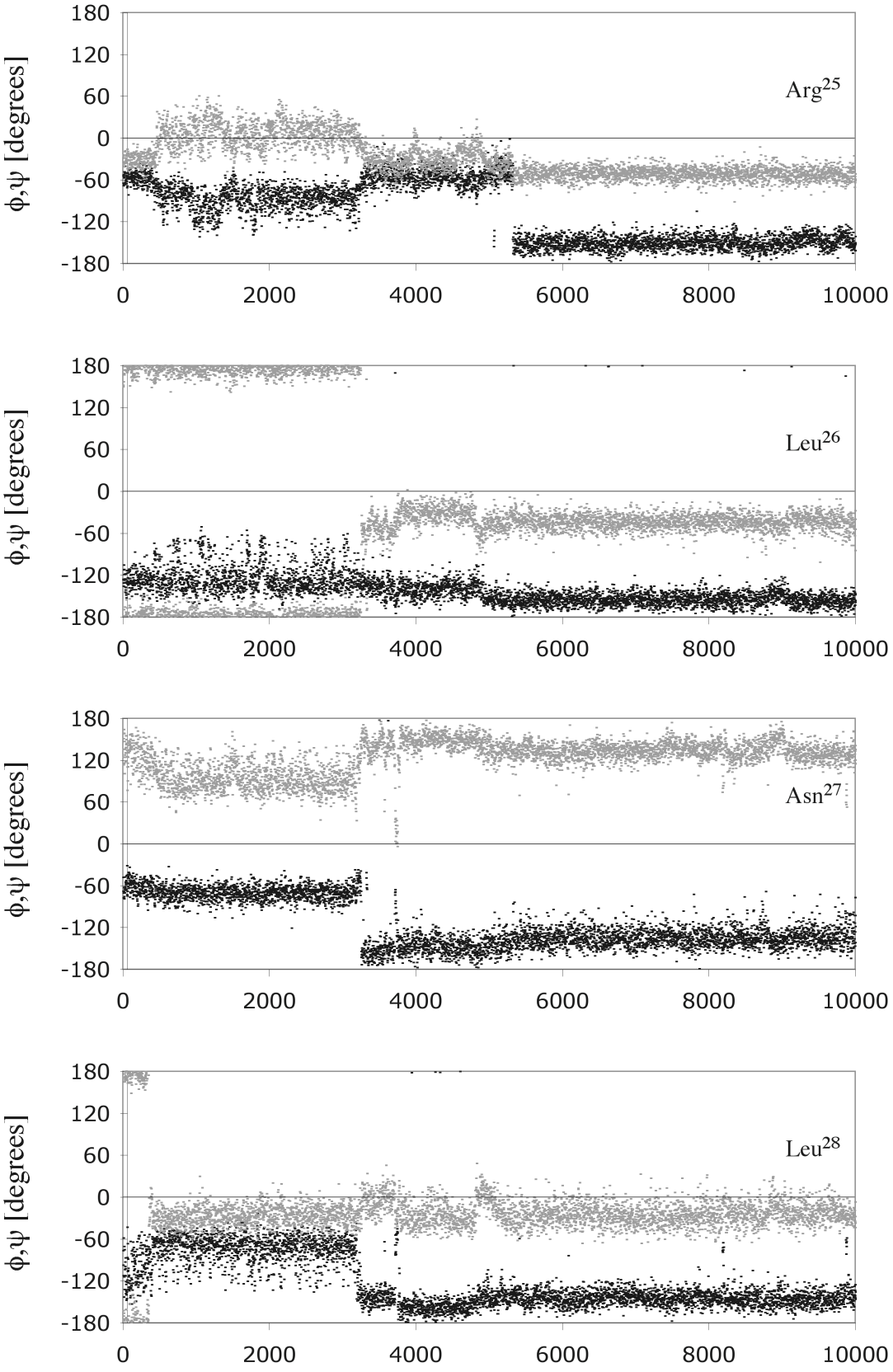

ϕ,ψ [degrees]
180
120
60
0
-60
-120
-180
0
2000
4000
6000
8000
10000
Trp5
Asn6
Arg7
Lys8

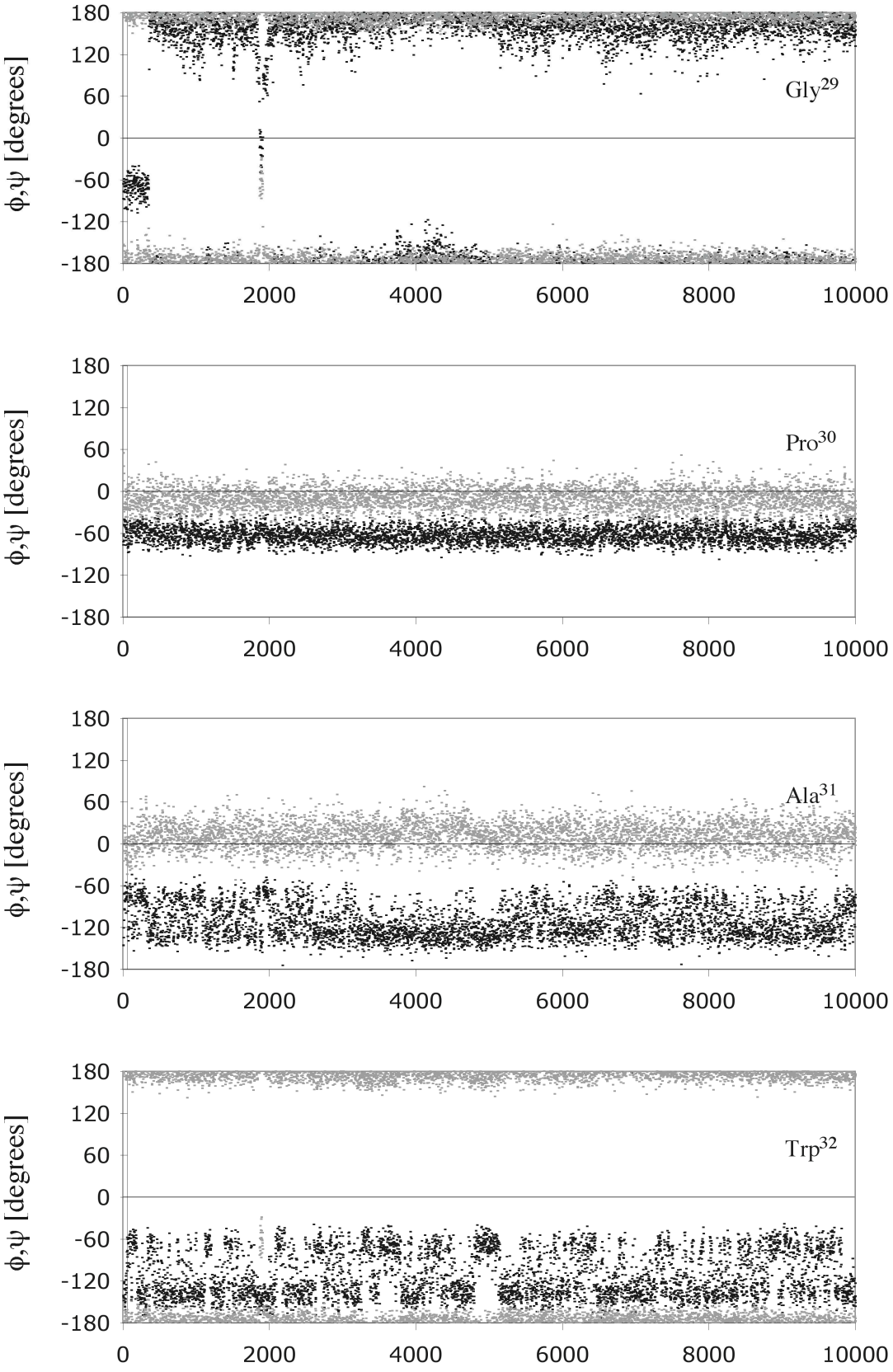

180
120
60
0
-60
-120
-180
ϕ,ψ [degrees]
0
2000
4000
6000
8000
10000
HYP9
Thr10
Lys11
Phe12

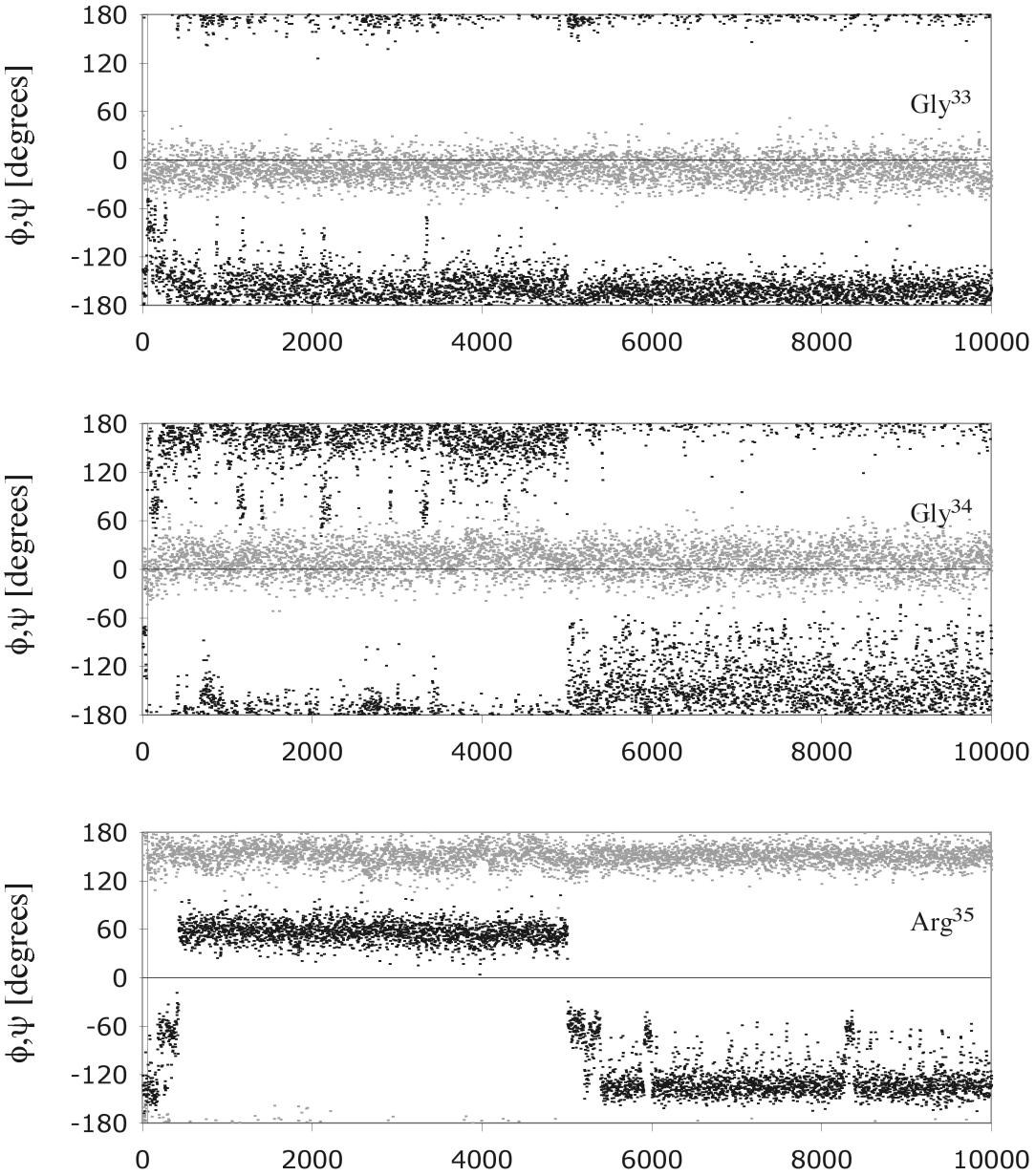

ϕ,ψ [degrees]
180
120
60
0
-60
-120
-180
0
2000
4000
6000
8000
10000
Hyp13
Ile14
Hyp15
Ser16


ϕ,ψ [degrees]
180
120
60
0
-60
-120
-180
0
2000
4000
6000
8000
10000
Hyp17
Asn18
Hyp19
Arg20


ϕ,ψ [degrees]
180
120
60
0
-60
-120
-180
0
2000
4000
6000
8000
10000
Asp21
Lys22
Trp23
Cys24


ϕ,ψ [degrees]
180
120
60
0
-60
-120
-180
0
2000
4000
6000
8000
10000
Arg25
Leu26
Asn27
Leu28


ϕ,ψ [degrees]
180
120
60
0
-60
-120
-180
0
2000
4000
6000
8000
10000
Gly29
Pro30
Ala31
Trp32


ϕ,ψ [degrees]
180
120
60
0
-60
-120
-180
0
2000
4000
6000
8000
10000
Gly33
Gly34
Arg35